LIBRARY
TECHNICAL REPORT SECTION
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

NPS55AA72051A

NAVAL POSTGRADUATE SCHOOL

Monterey, California

EVALUATION AND INNOVATION IN THE NAVY'S

PERSONNEL RESEARCH LABORATORIES

II. Development, Validation, and Trial Application of a Computer Program to Facilitate Judgmental Appraisal

by

James K. Arima

Richard W. Mister

May 1972

Approved for public release; distribution unlimited.

FEDDOCS
D 208.14/2:
NPS-55AA72051A

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral A. S. Goodfellow
Superintendent

M. U. Clauser
Provost

ABSTRACT:

The previous study in this series showed that evaluation of R&D activities rests eventually on qualitative judgments. The purpose of this study was to develop, validate, and test apply a procedure for obtaining qualitative judgments economically and efficiently. The Ford procedure for scaling partially ordered sets of rankings was programmed and validated using an abstract judgmental task with an extrinsic criterion. It was given a trial application requiring the ordering on merit of current personnel research projects. Both validation and trial application results were highly satisfactory. It was concluded that the Ford procedure could be used to obtain scaled qualitative judgments in a wide variety of settings with accuracy, efficiency, and economy. Flow charts, data setup, and the complete computer program are given.

This research was supported in part by the Personnel Research Division, Bureau of Naval Personnel, through Project Order No. 1-0001, Naval Personnel Research and Development Laboratory.

Prepared by:

PREFACE

This report is the second in a research project between the sponsoring activity, the Personnel Research Division of the Bureau of Naval Personnel, and the Naval Postgraduate School. The study was performed under the auspices of Capt. G. F. Britner, Division Director, and Mr. A. A. Sjoholm, Technical Director, Personnel Research Division.

We would like to express our thanks to Dr. Frank M. Andrews, Survey Research Center, Institute for Social Research, University of Michigan, for providing a copy of the Michigan Ford Program on which much of this work was based.

Portions of this work were done for a master's thesis in operations research by the junior author under the direction of the senior author.

Various aspects of this work were presented at the Research and Development Working Group, 28th Military Operations Research Symposium, Ft. Lee, Va., in November 1971, and at the XIXth International Meeting of The Institute for Management Sciences, Houston, Texas, in April 1972. The distribution list reflects the requests for this paper as a result of these presentations. It is hoped that recipients of this report will find it useful in the many different contexts of research indicated by their addresses and positions.

Evaluation and Innovation in the Navy's Personnel Research Laboratories, II. Development, Validation, and Trial Application of a Computer Program to Facilitate Judgmental Appraisal, by James K. Arima and Richard W. Mister.

BRIEF

The previous study in this series showed that there are no generally applicable, hard measures of the effectiveness of an R&D laboratory's activities and operations. The basis for determining effectiveness of a laboratory eventually narrows down to the judgments of persons who, for various reasons, are deemed qualified to make such judgments.

This being the case, it follows that the evaluation process can be improved by developing practical methods for obtaining and processing judgments that are simple to apply, broadly applicable, and faithfully reflect the contribution of each judge. Ideally, the results should be expressed quantitatively to permit their use in conjunction with other statistical and mathematical tools.

To have these characteristics, a method should permit an individual judge, faced with a set of alternatives to "prioritize", to rate only those with which he is familiar, to set his own measurement scale, and to make use of ties when he sees no difference between alternatives. The Ford procedure permits a judge to behave in this manner. It was originally programmed for computer application by the Survey Research Center, University of Michigan. The program was obtained and adapted for use on the computing facilities of the Naval Postgraduate School (NPS) which uses an IBM 360/67 system. The program, along with explanatory instructions, is reproduced in this report.

To prove the Ford program was broadly applicable and effective, a validation test was conducted using an abstract, vague, rating task for which there was--unknown to the judges--an independent set of quantitative "truth" data for comparison. Next, a trial application of the program was made in which Navy officers rated current personnel research projects as to the advisability of retaining and pursuing them in the R&D program. Finally, the Ford procedure was used in a real-life situation to analyze student ratings of courses in the NPS operations research program. The Ford rating procedure and NPS computer program were highly satisfactory in all of these test applications.

It was concluded that a simple, effective, and broadly useful procedure for obtaining and scaling the evaluative opinions of judges had been developed, tested, and applied. The suggestion was made to use the procedures to analyze project selection in the Navy's personnel research laboratories, since it is widely recognized that, for a laboratory to be effective, it must be working on the right programs at the right time.

TABLE OF CONTENTS

APPENDIXES

LIST OF TABLES

LIST OF FIGURES

I. PURPOSE AND SCOPE

The previous study in this series (Arima, 1971) discussed various factors associated with the effectiveness of Federal in-house laboratories. The problem of evaluating the effectiveness of a specific laboratory, such as the Navy's personnel research laboratories, was of special interest. Approaches to this evaluation problem seemed ultimately to require a qualitative assessment of a laboratory's effectiveness or some aspect of its operations by knowledgeable individuals. Accordingly, one specific problem identified as a result of the preliminary study was to develop and test a method for obtaining and analyzing such assessments from qualified judges in an economic, convenient, and effective manner. This report addresses itself to this problem.

The approach taken to solve the problem, explicated in the pages that follow, was: (1) Adapt Ford's (1957) procedure, as programmed by Pelz and Andrews (1966), for creating numerical rankings from a set of incomplete comparisons of objects by a group of judges to operate on the Naval Postgraduate School's IBM 360/67 system, (2) validate the procedures using an arbitrary task with an extrinsic criterion measure, and (3) test the feasibility of using the procedures to obtain an ordered set of qualitative judgments on an R&D problem appropriate to the environment and mission of the Navy's personnel research laboratories.

II. THE FORD PROCEDURE

A. FLEXIBILITY OF PROCEDURES.

There are three characteristics of Ford's procedure that make it especially appropriate for obtaining judgments on several alternatives or items from a diverse group of judges. First, a judge or rater adjudicates only those items that he feels competent to judge. Second, he can make his judgments as coarse or as fine as he desires because there is no restriction on how many judgmental categories he must use. And third, there is no requirement for a fixed distribution of items among the categories, except that, collectively over judges, no more than one third of all items being rated should be in any one category. A judge, for example, might decide to judge only half of a pool of items using three categories--high, medium and low.

The ease of this method can be compared with other frequently used methods that may require one or more of the following restrictions: all items must be ranked with no ties, each items is to be compared with every other item with no indeterminate category permitted, an equal number of items must be placed in each rating category, and so forth. Such restrictions are usually imposed because of statistical considerations in the analytical procedures. Unfortunately, persons who are unfamiliar with the statistical considerations are alienated against the results of the procedures because, while serving as judges, they had to make too many arbitrary decisions in which they felt no confidence. A more serious consequence of such procedures is the fact that a large amount of noise might be added to the

judgments so that the "signal" present in the judgments cannot be discriminated. Moreover, some of the techniques, such as paired comparisons, are excessively demanding of a judge's time. Thus, the statistical rigor is offset by serious negative consequences of the procedures involved.

At this time, it should be noted that the procedures being developed here are not the same as those designed to achieve a consensus or decision among a group of judges, such as some applications of the Delphi technique. These procedures tend to be used when the number of alternatives and judges are few, when any of the alternatives are reasonable choices, and when the problem is one of reaching consensus rather than evaluating the relative merit of the alternatives. The procedures tend to disregard the contribution of the individual and depend on devious group processes and feedback to eliminate, eventually, any individuality not consonant with the prevailing group trend. It should be pointed out that there is no way to determine to what extent the final decision is based on the relative merits of the items entering into the decision and on the group processes employed in arriving at a consensus. The procedures being developed here, on the other hand, produce a composite judgment that reflects the contribution of each judge according to the proportionate number of judgments he makes. The results of the procedure do not, however, produce a clear-cut decision or unanimity of opinion. Other factors and other methods must be employed for the decision-making process using the composite judgments as a data base. Bartee (1971), for example, suggests a linear program-

ming approach with zero-one variables. In many cases, however, the scaled alternatives might be an end in themselves with actions taking on priorities according to their scaled values.

B. DETAILS OF THE FORD PROCEDURE

The Ford procedure is based on forming a win-loss matrix, $A = (a_{ij})$, where a_{ij} represents the number of times object i is preferred over object j by the judges, and $a_{ii} = 0$. Moreover, all ties and nonjudged items do not enter the matrix for any one judge since a win-loss determination has not been made. Thus, each judge contributes to the composite judgment only those pairwise instances in which he has preferred one alternative over another. The Ford procedure then determines a weight, w_i, for each item. These weights are interpreted as odds in the sense that the probability of item i being preferred to item j in any comparison is taken to be $w_i/(w_i + w_j)$. These probabilities could then be used to compute matrix A. The set of these weights is the maximum likelihood of obtaining the original matrix, A. The weights are obtained by solving iteratively the equation

$$w_i^{n+1} = \frac{\sum_j a_{ij}}{\sum_j \frac{a_{ij} + a_{ji}}{w_i^n + w_j^n}} \qquad (1)$$

where a_{ij} = number of times object i was preferred to object j; a_{ji} = number of times object j was preferred to object i; w_i^n = weight assigned to object i on the n^{th} iteration; and w_j^n = weight assigned to object j on the n^{th} iteration. The weights are win

percentages on the first iteration. The iteration stops in the computer program when a predetermined convergence criterion is reached or a predetermined number of iterations has been completed.

There was one assumption in Ford's procedure that made it difficult to apply in practice. This was a partition assumption which stated that in any partition of the win-loss matrix into two nonempty subsets, some item in each subset had to be preferred at least once to some item in the other subset. That is, the initial w_i and w_j could not be 1 and 0 in equation (1). This rule would be broken in the case of universally high and universally low alternatives and in any subset where all judgments are in one direction. Pelz and Andrews (1966) solved this problem by first removing universally high and low items from the win-loss matrix before computing the weights and by adding a very small constant, .00001, to each of the remaining entries in the matrix. These procedures permitted them to program Ford's procedures for computer processing of judgments involving 130 judges and 130 items. Accordingly, the Pelz and Andrews program was used as a starting point for adapting Ford's procedure to the Naval Postgraduate School's IBM 360/67 system. The program as adapted for the IBM 360/67 system will hereafter be referred to as the Ford program.

C. THE FORD PROGRAM

A flow-chart of the program is included at Appendix I. The data assembly for input to the program is shown in Appendix II. The program, itself, with explanatory comments is reproduced at Appendix III.

Two decisions are required by the person using the program. First, he must specify the convergence criterion for the iterative determination of the weights. This report uses .005. That is, when the weights do not change by that amount in successive iterations, a satisfactory stabilization of the weights is accepted. Second, the user must specify how many iterations are to be conducted in the event the convergence criterion is not reached. This reports uses 50. As will be shown, the rank ordering of the items, as determined from their weights, stabilizes rapidly. Accordingly, even if the convergence criterion is not met, the rank ordering is acceptable. When the convergence criterion is met, the weights can be used as an interval scaling of the judged items.

The program operates in three subroutines or cores. The first core assigns an ID number (hereafter called "assigned ID number") to each rated alternative as it is read into the computer and them computes how many comparisons are to be made between pairs of alternatives, excluding ties.

The second core forms the win-loss matrix, eliminates universal highs and lows, assigns the small constant to each cell, and then computes the initial weights.

The third core performs the iterations until the weights stabilize or until the specified number of iterations have been run. The results are printed out showing a list of judges and the number of comparisons made. The output gives a mapping of the assigned ID numbers to the original numbers used for input of the variables. The win-loss

matrix is shown by assigned ID number. Finally, there is a printout of the weights by iterations and a list of final weights shown by assigned ID number and giving the corresponding original ID number.

III. VALIDATION OF THE FORD PROGRAM

A. THE VALIDATION PROBLEM

Pelz and Andrews (1966) showed some comparisons of the Ford procedure with alternative methods for scaling partially ordered judgments. Having shown the computational advantages of the Ford procedure, they then demonstrated its utility in their evaluation of scientists in organizations. They did this by having laboratory directors rate their scientists as to their excellence in scientific research using the Ford procedure. These ratings were then scaled and used as the criterion variable in their studies. It should be noted, however, that the validity of these ratings was not established in a psychometric sense (American Psychological Association, 1954), other than that of face validity. That is, they were not subjected to a critical comparison against some outside criterion.

Among the other forms of validity--concurrent, predictive, and construct--concurrent validity of the scaled judgments would be of most interest when the judgments are to be used as a criterion measure, dependent variable, objective function, or, in general, a measure of effectiveness. That is, we would like to know how well the judgments represent the true state of the world that they are presumed to repre-

sent. This is particularly true when, as in the case of the Ford procedure, judgments which are ordinal in nature are mapped to the system of real numbers and used as a cardinal measure. For the application made by Pelz and Andrews, we would like to know how accurately the scaled ratings represent the true effectiveness of the rated scientists. Stated in this form, the difficulty or impossibility of assessing the concurrent validity of the scaled ratings becomes readily apparent: judgments of this type are used because there is no other acceptable measure of the variable in which interest lies.

In view of the foregoing, it follows that an existing, scaled variable is needed to validate the Ford program. In its simplest form, validation might take on the paradigm of a psychophysical experiemnt. For example, a set of standard weights might be presented to judges with the task of rating the relative heaviness of the weights. There would be little interest in such a test of the Ford procedure, since it would be a straightforward evaluation of a numerical estimation function as the size of the weights vary. In a validation of the Ford procedure, interest lies in the nature of the underlying quality of pairs of objects as they are judged and what the relationship is of the perceived quality to the decisions of the judges. This distinction in emphasis is elaborated in detail by Krantz (1972). The test in a psychophysical paradigm might be more relevant, for example, if the judges had to rate the weights of objects differing considerably in size and mass. Thus, an ideal validation of the Ford procedure would

take place if judges were to rate items according to an abstract or vague variable for which there is, unknown to them, a corresponding quantitative, objective variable that could serve as a criterion measure. Unfortunately, the more vague or abstract a judging task becomes, the more difficult it is to find a criterion variable that is also not equally vague. Accordingly, validation of the Ford procedure with a challenging task will be less than rigorous and any discrepancy of the resulting scaled judgments from the criterion values may be due to several factors which will not be independently assessable. These include the difficulty of the judgmental task, the capability of the judges, the reliability of the criterion variable, and the efficiency of the Ford program. The validation, then, will be clinical, and hopefully diagnostic, while attempting to be rigorous.

B. METHOD

1. Stimulus Materials.

Fortunately, there is a situation that compares favorably with the ideal validation paradigm mentioned above. It has been found that such abstract characteristics or qualities of words as their familiarity, meaningfulness, and associational richness are closely related to the frequency with which they appear in the English language (Broadbent, 1967; Ekstrand, Wallace, & Underwood, 1966; Underwood, 1966). Fortunately, too, the frequency of 30,000 words has been cataloged in what has become known as the Thorndike and Lorge (1944) word count. Now, it can be assumed that most individuals are not consciously aware of the

fact that familiarity, say, of English words depends on their frequency. Accordingly, it should be possible to ask judges to rate a list of selected words from the Thorndike and Lorge word count for their familiarity to persons in general and compare the Ford-scaled ratings with the Thorndike and Lorge word count, thus completing the validation.

Rather than selecting words directly from the Thorndike and Lorge word count, an intermediate procedure was inserted to provide some structure to the judging task and to make possible four replications of the judging procedure. The words were actually taken from the category norms for verbal items compiled by Battig and Montague (1969). Their norms are based on the primacy and frequency with which students at two large universities provided verbal associations for 56 different verbal categories, such as a precious stone, a unit of time, and so forth. Four of these categories were chosen from which to select words based on the fact that there was a correlation of .90 or greater between the two universities and that there was a long enough list of associations from which selections could be made, graded for their frequency in the Thorndike and Lorge count. The categories selected, which will hereafter be referred to only by the Roman numeral designation given below, were:

I. A kind of cloth ($r = .988$)

II. A kitchen utensil ($r = .987$)

III. A substance for flavoring food ($r = .977$)

IV. A disease ($r = .906$)

The correlations shown are those between the two university groups, and are based on the rank position occupied by the words within a category based on their frequency of mention.

The selection of specific words from the categories was made by reference to the Thorndike and Lorge word count using the following guidelines, which could be applied only approximately. Twelve words were chosen from each category to make a test list. The 12 words were further divided into approximately four groups with at least a 5 to 10 percent difference in frequency of occurrence between each group, based on the Thorndike and Lorge general (G) count. Between items in each group, there was a 1 to 3 percent difference in the frequency of occurrence. Where there were ties in the general count, the other counts (T, L, and S) given in the word count were used to break the ties. Thus, there was a fairly reliable clustering of words into four frequency ranges within each list and a less reliable ranking within the frequency ranges. The lists are shown in Table I. Each category provided an independent relication for validation.

2. Subjects

Twenty male and female Naval Postgraduate School students ranging in age from 24 to 37 years with comparable levels of education served in the validation experiment. Each subject was used twice, and 10 subjects were assigned at random to each of the four categories.

3. Procedure

Each word list was reproduced in random order on a sheet of paper. The subjects were told to make an ordinal ranking of the words as to what they believed their relative familiarity was to all people in general. They were further instructed to judge only those objects which they could rank with confidence, make use of as many ordinal ranks as they deemed necessary, and to place as many objects in each rank as they desired. By way of guidance, they were instructed to select the number of ordinal ranks they were willing to use first and then to write the number of the rank beside the objects they chose to rank. They were also advised to give first impressions and work rapidly.

C. RESULTS

The orderings made by the subjects and processed by the Ford program are shown in Table 2, along with the Spearman rank correlation (rho) between the Thorndike-Lorge and Ford program orderings. The results will be examined in detail only for category I.

Table 3 shows the win-loss matrix for category I. The rows (i) are arranged in the sequence, from top to bottom, according to their assigned ID numbers. When one reads across the table horizontally, he is reading the number of times the row item was preferred to any column item and the sum in the rightmost column shows how many times the row item "won." When one reads down the columns vertically, he is reading the number of times the column item lost to the row item, and the sum

TABLE 1

VALIDATION TEST LISTS WITH WORDS PRESENTED IN

THORNDIKE-LORGE RANK ORDER WITHIN CATEGORIES

CATEGORY I	CATEGORY II	CATEGORY III	CATEGORY IV
1. cotton	1. cup	1. salt	1. cold
2. felt	2. bowl	2. sugar	2. rheumatism
3. wool	3. knife	3. sage	3. typhoid
4. lace	4. fork	4. ginger	4. cancer
5. velvet	5. refrigerator	5. vinegar	5. smallpox
6. canvas	6. saucer	6. cloves	6. cholera
7. muslin	7. sieve	7. mustard	7. measles
8. pique'	8. skillet	8. cinnamon	8. rheumatic fever
9. rayon	9. ladle	9. nutmeg	9. syphilis
10. corduroy	10. scraper	10. thyme	10. diabetes
11. denim	11. toaster	11. basil	11. dysentery
12. batiste	12. cleaver	12. cayenne	12. peritonitis

TABLE 2

FORD PROGRAM RANK ORDERING OF WORDS WITHIN CATEGORIES

(Criterion rank numbers and the Pearson rank order correlation between the computer and criterion rank orders are shown.)

CATEGORY I (rho = .521)

1.	cotton
3.	wool
4.	lace
6.	canvas
10.	corduroy
5.	velvet
11.	denim
9.	rayon
2.	felt
8.	pique'
7.	muslin
12.	batiste

CATEGORY II (rho = .598)

4.	fork
6.	saucer
3.	knife
1.	cup
2.	bowl
8.	skillet
11.	toaster
5.	refrigerator
12.	cleaver
9.	ladle
10.	scraper
7.	sieve

CATEGORY III (rho = .687)

1.	salt
2.	sugar
7.	mustard
5.	vinegar
6.	cloves
8.	cinnamon
9.	nutmeg
4.	ginger
3.	sage
11.	basil
12.	cayenne
10.	thyme

CATEGORY IV (rho = .460)

1.	cold
4.	cancer
7.	measles
9.	syphilis
10.	diabetes
2.	rheumatism
5.	smallpox
6.	cholera
3.	typhoid
11.	dysentery
8.	rheumatic fever
12.	peritonitis

TABLE 3

WIN-LOSS MATRIX FOR CATEGORY I

(Computer-assigned ID numbers are shown in parentheses)

i	j (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	$\sum a_{ij}$
(1) Batiste	0	0	0	0	1	1	1	1	1	0	1	1	7
(2) Lace	6	0	5	2	5	3	7	9	5	7	5	1	55
(3) Pique'	2	0	0	1	1	1	1	1	1	0	1	1	10
(4) Cotton	5	6	5	0	9	7	10	10	7	8	7	4	78
(5) Felt	5	1	5	0	0	2	3	7	2	4	2	1	32
(6) Canvas	5	4	5	0	7	0	6	9	5	7	5	0	53
(7) Denim	5	1	5	0	5	2	0	6	4	2	4	1	35
(8) Muslin	4	0	4	0	2	1	0	0	2	1	1	1	16
(9) Velvet	5	2	5	0	4	2	5	8	0	5	2	1	39
(10) Rayon	6	2	5	1	4	1	4	8	2	0	3	1	37
(11) Corduory	5	2	5	0	5	1	6	9	2	5	0	1	41
(12) Wool	5	6	5	1	8	5	7	9	6	8	5	0	65
Sum (losses)	53	24	49	5	51	26	50	77	37	47	36	13	468
Win Percent	11.7	69.6	16.9	94.0	38.6	67.1	41.2	17.2	51.3	44.0	53.2	83.3	

at the bottom of the columns show the frequency of losses. Within the matrix, any entry shows how many times a comparison was made between the two items involved. For example, the maximum number of 10 comparisons was only made between cotton and denim and cotton and muslin. In the matrix notation, these would be $a_{4,7}$ and $a_{4,8}$. The win percents which would be used as the initial weights in equation (1) are shown below the column sums. A comparison of the rankings which would be made on the basis of the Thorndike-Lorge Count, the Ford program scaling, and the win percent are shown in Table 4. A graph showing how the weights change per iteration is presented in Figure 1.

The observed rank correlation of .521 between the Thorndike-Lorge and category I rankings is not as high as one would like. An examination of the rankings showed a great discrepancy for the word, felt. Two good reasons can be given for this discrepancy with the benefit of retrospect. First, it was found that "felt" in the Thorndike-Lorge count includes the past tense of "feel", which would account for its high position in the word count. The cloth, felt, is included also. The subjects were, of course, ranking the latter use of the word. Second, the Thorndike-Lorge count was published in 1944 and the use of felt has diminished greatly since then so that it is not as familiar to a new generation of persons. Recomputation of the correlation for category I with felt removed resulted in a rho of .788.

Similarly, rho of .460 was disappointing for category IV (Table 2). Inspection of the differences in rankings showed typhoid and syphylis occupying diametrically opposite positions in the two rankings

TABLE 4

COMPARATIVE RANK ORDERING OF CATEGORY I ITEMS

Thorndike-Lorge	Ford Program	Win Percent
1. Cotton	1	1
2. Felt	9	9
3. Wool	2	2
4. Lace	3	3
5. Velvet	6	6
6. Canvas	4	4
7. Muslin	11	10
8. Pique'	10	11
9. Rayon	8	7
10. Corduroy	5	5
11. Denim	7	8
12. Batiste	12	12

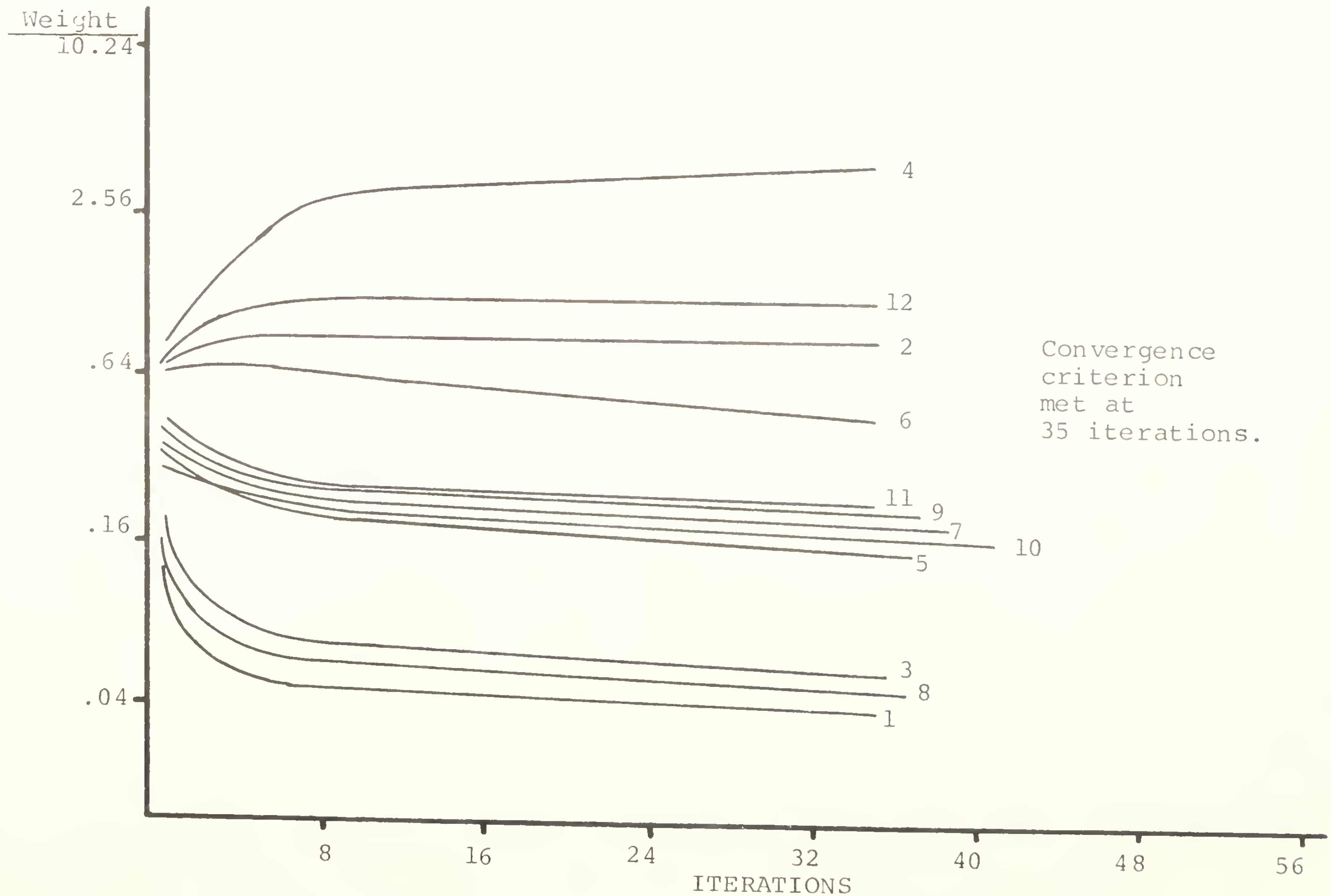


Figure 1. Weight Change per Iteration. Category I.
Assigned ID Number Used.

(Table 1). The differences could again be accounted for by changing trends in the incidence of the diseases and the openness with which syphylis is mentioned today compared with 1944. Moreover, the military personnel who served as subjects would be more sensitive to syphylis as a disease than the population at large owing to the emphasis given venereal disease prevention in the military services. With the differences in the observed ranks halved for the two diseases, rho for category IV was increased to .585. With these two changes, each of the four obtained correlation coefficients was found to be significantly different from a hypothesized rho of zero by a 2-tailed, $\underline{t}$ test at the .05 level.

Table 4 also suggests that the win percent calculated from the win-loss matrix is closely related to the final ordinal rankings of the items. In consonance with this observation, it was found that rank order stability was reached after the first iteration for categories I, III, and IV and after the third iteration for category II. Category I converged in 35 iterations and category III, in 16. No convergence was reached for categories II and IV after 50 iterations. Four objects in category III were rated as universal highs and were removed prior to computation of weights.

D. DISCUSSION AND SUMMARY

To recapitulate, the validation procedure used 20 individuals who were assigned in groups of 10 to four tasks requiring them to make ordinal judgments that were made purposefully difficult. The results

showed that in all four cases the judgments made by the group were significantly related to the criterion, that ordinal rankings of the judged items were made quickly and efficiently, and that in two of the four tasks, the numerical scaling of the ranked items had converged to a stable position. The magnitude of the corrected correlation coefficients showed that approximately 30 to 60 percent of the total variance was accounted for in the correspondence between judgments and the criterion. This is considered excellent in view of the many factors that operated to attenuate the correlation coefficients. First, as mentioned above, the criterion was based on old information. Moreover, the criterion was based on a word count made entirely from printed materials, whereas the task given the judges implied familiarity of the words based on all contexts. Too, the Thorndike-Lorge word count used all meanings of the words--e.g., ginger as a seasoning and a girl's name, sage as a seasoning and a wise man--whereas their familiarity was judged in the specific category specified. Additionally, the crucial assumption that made this validation possible--that familiarity with verbal materials is related to their frequency of occurrence in the language--is in itself not a perfect relationship. Another factor that was no doubt a severe constraint on the magnitude of the correlations was the way the words were chosen for the lists. That is, there was a very minute difference in the frequency count of some words within their selection bands. In fact, two words in one of the middle bands and all four words in the bottom band of the category I list were tied in frequency in the Thorndike and Lorge general count. This was done

to ensure that there would be a large number of ties in the rankings of the judges in order to make a thorough test of the Ford program. Considering the total impact of these attenuating factors, the obtained correlation coefficients are very high and provide strong evidence for the efficiency of the Ford ranking procedure and the Pelz and Andrews computer program as adapted for the Naval Postgraduate School's IBM 360/67 system.

IV. TRIAL APPLICATION OF THE FORD PROGRAM

A. PROBLEM SELECTION

It has been shown that the Ford program is effective in taking ratings of judges with respect to an abstract, qualitative dimension and scaling them. The next and final step in this project is to determine whether the procedures can be efficiently and effectively applied to a practical problem. If the former test can be considered a validation of the program, the next step could be called a trial application of the program.

It would be desirable to have the trial application duplicate in detail a planned or proposed actual use of the Ford program. Now, it was emphasized in the previous report (Arima, 1971) that proper project selection was a crucial component of successful laboratory management. Dr. Donald F. Hornig, then director of the Office of Science and Technology in the Executive Office of the President, was quoted as saying that one of the most critical questions in the effective utilization of Federal laboratories was "The choice of problems,

their significance, and the feasibility of finding solutions through research and development . . . ." (Subcommittee, 1968; p. 9). One way to improve project selection might be to examine current projects for their significance using representatives of sponsor and using agencies, and to examine the feasibility of finding solutions through research and development by having in-house scientific/technical personnel evaluate current projects from this standpoint. This line of reasoning led the trial application of the Ford program to the problem of evaluating the significance of current programs.

B. METHOD

1. Stimulus Materials.

As part of the review of in-house laboratories being conducted by the Director of Defense Research and Engineering, the Director of Navy Laboratories by letter dated 25 March 1971 requested various activities within the Navy to document significant contributions and accomplishments by their in-house laboratories. Using the material prepared in response to this request by the Personnel Research Division, Bureau of Naval Personnel, for the Navy's personnel research laboratories, 10 programs were selected at random as items to be rated for this trial application of the Ford program. The project descriptions given in the report were edited and condensed, in some cases, and appear in Appendix IV. A listing of the programs chosen is shown below. The numbers and/or the short title (in parentheses) given in the listing will hereafter be used to reference and identify the programs. The programs were:

(1) Improved Enlisted Personnel Distribution and Management (Personnel Distribution)

(2) Ship Manning Requirements Techniques (Manning Requirements)

(3) Evaluation of Standards for Navy Reenlistment (Reenlistment Standards)

(4) Development of Navy Military Personnel Costing Techniques for Use in Determining Cost Implications Associated with Changes in Reenlistment Rates (Reenlistment Costing)

(5) Design of an Optimum Personnel Force Structure (Personnel Structure)

(6) Interest Measurement in Officer Selection (Officer Selection)

(7) Evaluation Survey of the Effectiveness of Submarine Sonar Operator Training (Sonar Training)

(8) Marginal Personnel/Minority Group Testing (Personnel Testing)

(9) Personnel Cost Research for Early Man/Machine Design Trade-Offs (Man-Machine Costs)

(10) LOFARGRAM Analysis Procedures (LOFARGRAM Analysis)

2. Subjects

The subjects were 10 Navy officer students attending the Naval Postgraduate School.

3. Procedure

The method was essentially identical to the validation procedures. Each subject was given a copy of the research programs (Appendix IV) and instructed to make an ordinal ranking of the items with respect to their

desirability and need for retention and further development as research programs within the Navy. As before, they were told to rank only those items which they could with confidence, to use as many ranks as they deemed necessary, and to place as many programs as they desired in any ranking category. They were advised to review the programs first and then decide on the number of ranking categories to use. Having done this, they wrote the number of the rank chosen beside the program description. Cards were keypunched from these data and run through the Ford program.

C. RESULTS

The rankings given the 10 programs by the 10 judges are shown in Table 5. The smallest number of programs ranked was four by judge number six. Another judge ranked 8 items, and the other eight judges ranked all programs. Of the latter, five judges used three categories; one used four; another five; and another, 10 categories. The number of comparisons made by each judge is shown in Table 6 for a total of 312 comparisons.

The win-loss matrix is shown in Table 7 with sums of wins (a_{ij}) and losses (a_{ji}) in the right and bottom margins, respectively. There were no universal highs or lows. Only 14 iterations were required to achieve stable weights at the .005 criterion. The program used 7.55 secs. of central processor unit time. Table 8 shows a summary of the results. The items are listed in the ordinal order of final ranks and show the number of comparisons in which each item was involved (sums of wins and losses), the win percent, and the final weights.

TABLE 5

RANKINGS OF TRIAL APPLICATION PROJECTS BY INDIVIDUAL JUDGES

Project	Judges 1	2	3	4	5	6	7	8	9	10
1. Personnel Distribution	2	2	1	1	1	1	2	1	2	1
2. Manning Requirements	2	1	1	5	2		1	2	1	7
3. Reenlistment Standards	1	1	3	7	3		1	1	1	5
4. Reenlistment Costing	1	2	2	4	3		1	2	3	2
5. Personnel Structure	2	1	1	6	2		3	1	4	4
6. Officer Selection	3	3	1	3	2	2	1	1	5	6
7. Sonar Training	1	2	3		3	2	4	3	3	3
8. Personnel Testing	3	1	2	8	2		4	3	2	10
9. Man-machine Costs	2	3	2	2	3		3	2	4	9
10. LOFARGRAM Analysis	2	2	3		1	1	4	3	4	8

TABLE 6

NUMBER OF COMPARISONS MADE BY EACH JUDGE IN

THE TRIAL APPLICATION TEST

Judge Number	Number of Comparisons
1	31
2	32
3	33
4	28
5	32
6	4
7	35
8	33
9	39
10	45
TOTAL	312

TABLE 7

WIN-LOSS MATRIX FOR THE TRIAL APPLICATION TEST

i	j: 1	2	3	4	5	6	7	8	9	10	$\sum a_{ij}$
1. Personnel Distribution	0	4	4	6	5	7	7	7	8	5	53
2. Manning Requirements	3	0	3	4	3	3	6	7	6	6	41
3. Reenlistment Standards	4	3	0	3	3	4	4	6	6	6	39
4. Reenlistment Costing	2	3	3	0	5	4	4	5	5	6	37
5. Personnel Structure	1	2	4	4	0	4	5	6	5	5	36
6. Officer Selection	1	3	3	4	2	0	4	5	5	4	31
7. Sonar Training	1	2	1	0	3	4	0	2	4	3	20
8. Personnel Testing	1	0	2	3	1	2	4	0	3	3	19
9. Man-machine Costs	0	1	2	1	1	3	3	5	0	3	19
10. LOFARGRAM Analysis	0	1	1	1	1	4	2	4	3	0	17
$\sum a_{ji}$	13	19	23	26	24	35	39	47	45	41	312

TABLE 8

SUMMARY RESULTS OF THE TRIAL APPLICATION TEST

	Project	Number of Comparisons	Win Percent	Final Weights
1.	Personnel Distribution	66	80.3	1.606
2.	Manning Requirements	60	68.3	.868
3.	Reenlistment Standards	62	64.4	.702
4.	Reenlistment Costing	63	60.2	.620
5.	Personnel Structure	60	60.0	.612
6.	Officer Selection	66	47.0	.338
7.	Sonar Training	59	35.4	.217
8.	Personnel Testing	66	28.8	.175
9.	Man-machine Costs	64	29.7	.174
10.	LOFARGRAM Analysis	58	29.3	.171

D. DISCUSSION AND SUMMARY

While the number of judges and the number of alternatives evaluated were small, consistent trends were evident. Except for one judge who only contributed four comparisons, all the other judges contributed from 28 to 45 comparisons, showing that any judge makes a significant contribution to the total number of judgments, even if he does not rank all items and uses few rank categories. Similarly, in spite of the freedom permitted the judges in choosing items to rate and the number of rating categories, the entries in Table 8 show that all items entered into a fairly uniform number of comparisons with a range from 58 to 66. Obviously, both of these distributions will depend on the sample of judges and the types and number of alternatives to be judged, but it is apparent from this trial that there will be a central tendency in the number of categories judges will choose to use and the number of alternatives a judge will adjudicate. Similarly, the alternatives will tend to attract a fairly uniform number of comparisons over a number of judges. Moreover, when the choices are difficult, there will probably not be any universal highs or lows, thanks to those who bet the long shots and the other who will give the lowest underdog a boost. The most important finding, however, was that the weights stabilized rapidly, indicating that a group of judges can achieve reasonable consensus in their composite judgment. Finally, the efficiency of the system was revealed by the very short computer time required for the scaling.

Five of the rated programs could be identified in the work plans of the two laboratories with some degree of certitude. From these

descriptions, the five were ranked according to FY1971 expenditures for each program, and a Pearson rank correlation coefficient was calculated with the ranks of the programs based on their weights obtained from the 10 judges. The obtained rho was .60, which suggests that there is a relationship between the amounts being invested in these research projects and the combined judgments of Naval officers who are representative of user elements of the Navy. This trend lends credence to the suggestion presented above, that the Ford program might well be used to analyze project selection based on the relationship between funding and user ratings, professional estimates of feasibility of finding solutions through research, and the resources actually being programmed for the projects.

V. ADDITIONAL APPLICATIONS

A. SITUATION

Concurrent with this study, an investigation was being made into the relative values of the major segments of the Naval Postgraduate School's operations research courses as seen by the student. One group of 54 graduating students in the operations analysis curriculum and another group of 15 graduating students in various management curricula had been asked to rank nine program segments in the operations research list of courses. The data lay unanalyzed because of the many ties (which were permitted) and because students had ranked different numbers of the program segments. (They could not rank courses they had not taken.)

B. RESULTS

The data were in a form that would be obtained in an application of the Ford procedure. Accordingly, they were run through the Ford program with a convergence criterion of .005. The criterion was reached in 14 iterations for the 54 operations analysis students and in 24 iterations for the management students. The result was a useful scaling of the items for the purposes that had motivated their collection.

C. COMMENTS

This application in a genuine research setting shows the utility of the Ford program. It confirms statements made above in the discussion of the trial application test that a consensus--in the form of weight convergence--is rapidly reached when knowledgeable judges rate clearly defined, real-world alternatives. One must conclude that the Ford program could be used to good advantage in the many, ever increasing, difficult, decision situations which are currently arising in which value judgments made by individuals are the major sources of data. It should be noted, too, that the data had been collected in a manner that was identical to an application of the Ford procedure. In this case, however, circumstances dictated that they be collected in this fashion. That is, the investigators felt that, to get a valid sampling of opinions, the individual judge had to be permitted to use the number of rating categories he desired (effectively accomplished by permitting multiple ties) and to refrain from adjudicating those items with which he was not familiar. That these elements should be characteristic of a good

scheme for collecting qualitative judgments was mentioned in the introductory portions of this study.

VI. SUMMARY AND CONCLUSIONS

The purpose of this study was to develop, validate, and test the feasibility of a procedure for obtaining qualitative judgments from individuals to be used in evaluating the effectiveness and operations of the Navy's in-house, personnel research laboratories. The Ford procedure for scaling partially ordered rankings, as programmed by Pelz and Andrews, was further programmed for the Naval Postgraduate School's IBM 360/67 system. The procedures and program were validated using an arbitrary, abstract task for which there was an extrinsic criterion and tested for feasibility in research evaluation using descriptions of actual program projects. In both cases, the results were highly satisfactory.

It can be concluded that the Ford procedure and present program can be used to obtain qualitative judgments with accuracy and efficiency. The utility of the program is limited only by the imagination and creativity of the user in devising appropriate rating schemes for his purpose. It should be a very useful tool for the many researchers who today are faced with analyzing "quality of life" variables for which conventional measurements do not exist.

REFERENCES

American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. Psychological Bulletin, 1954, 51, (2, Part 2).

Arima, James K. Evaluation and innovation in the Navy's personnel research laboratories, I: Preliminary considerations. Technical Report NPS55AA71071A, Naval Postgraduate School, July 1971.

Bartee, E. M. Problem solving with ordinal measurement. Management Science, 1971, 17, B-622-B-633.

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories. Journal of Experimental Psychology, 1969, 80, (3, Part 2).

Broadbent, D. E. Word-frequency effect and response bias. Psychological Review, 1967, 1-15.

Ekstrand, B. R., Wallace, W. P., & Underwood, B. J. A frequency theory of verbal discrimination learning. Psychological Review, 1966, 73, 566-578.

Ford, L. R., Jr. Solution of a ranking problem from comparisons. American Mathematical Monthly, 1957, 64, No. 8, Part II, 28-33.

Krantz, D. H. Measurement structures and psychological laws. Science, 1972, 175, 1427-1435.

Pelz, D. C., & Andrews, F. M. Scientists in organizations. New York: Wiley, 1966.

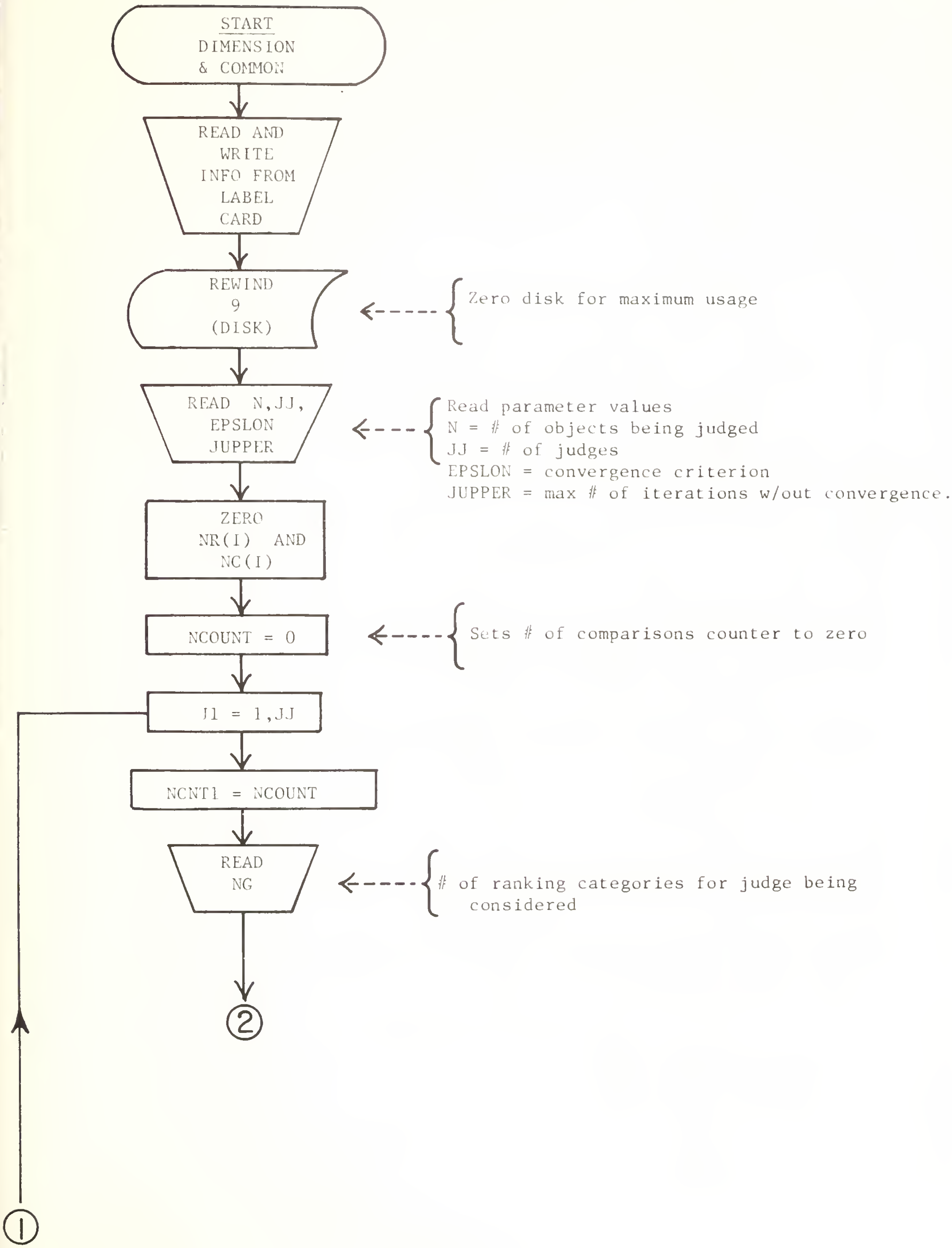
Subcommittee on Science, Research and Development; Committee on Science and Astronautics, US House of Representatives. Utilization of federal laboratories. Washington: Government Printing Office, 1968.

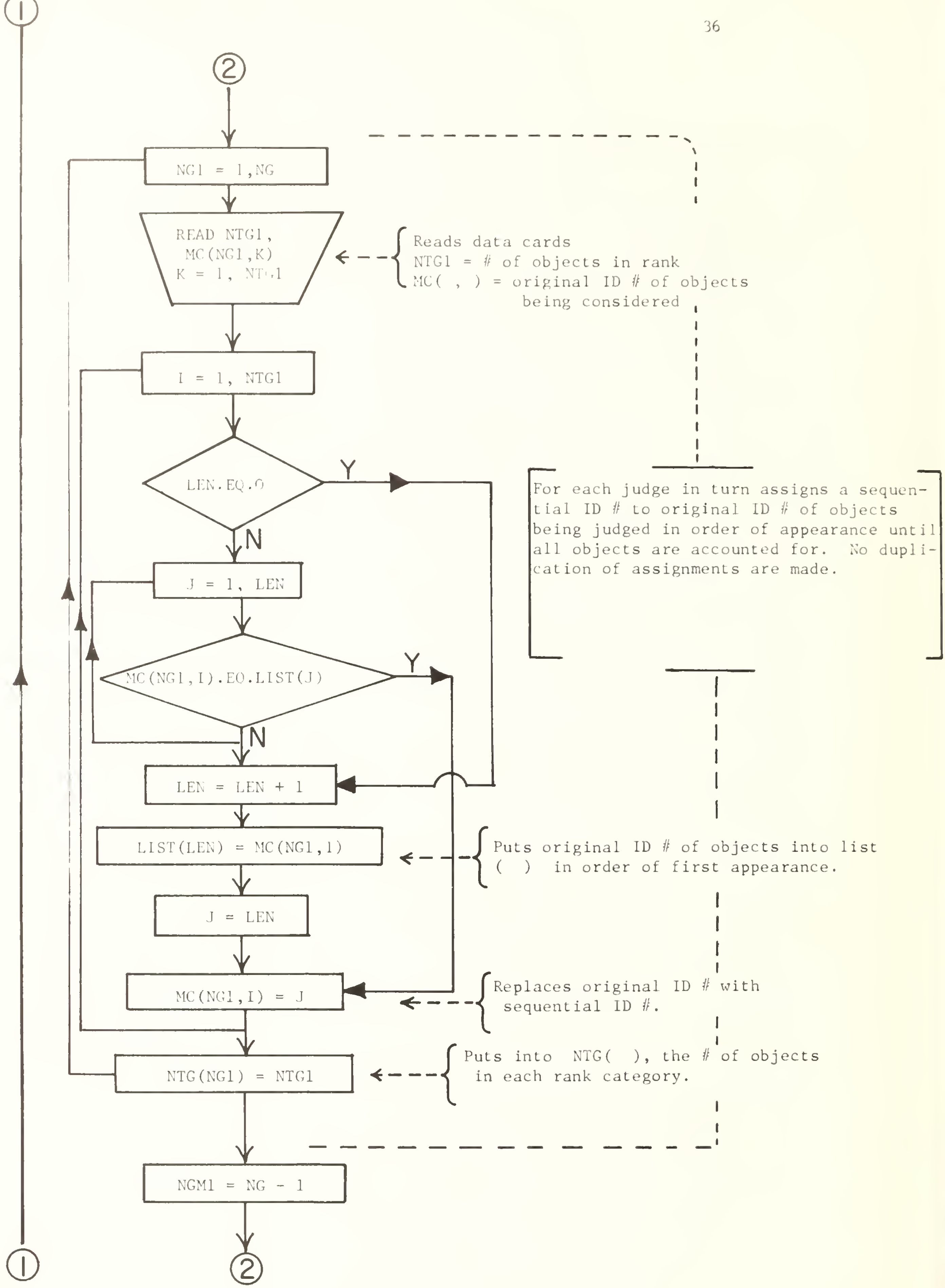
Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Columbia University, 1944.

Underwood, B. J. Experimental psychology (2nd ed.) New York: Appleton-Century-Crofts, 1966.

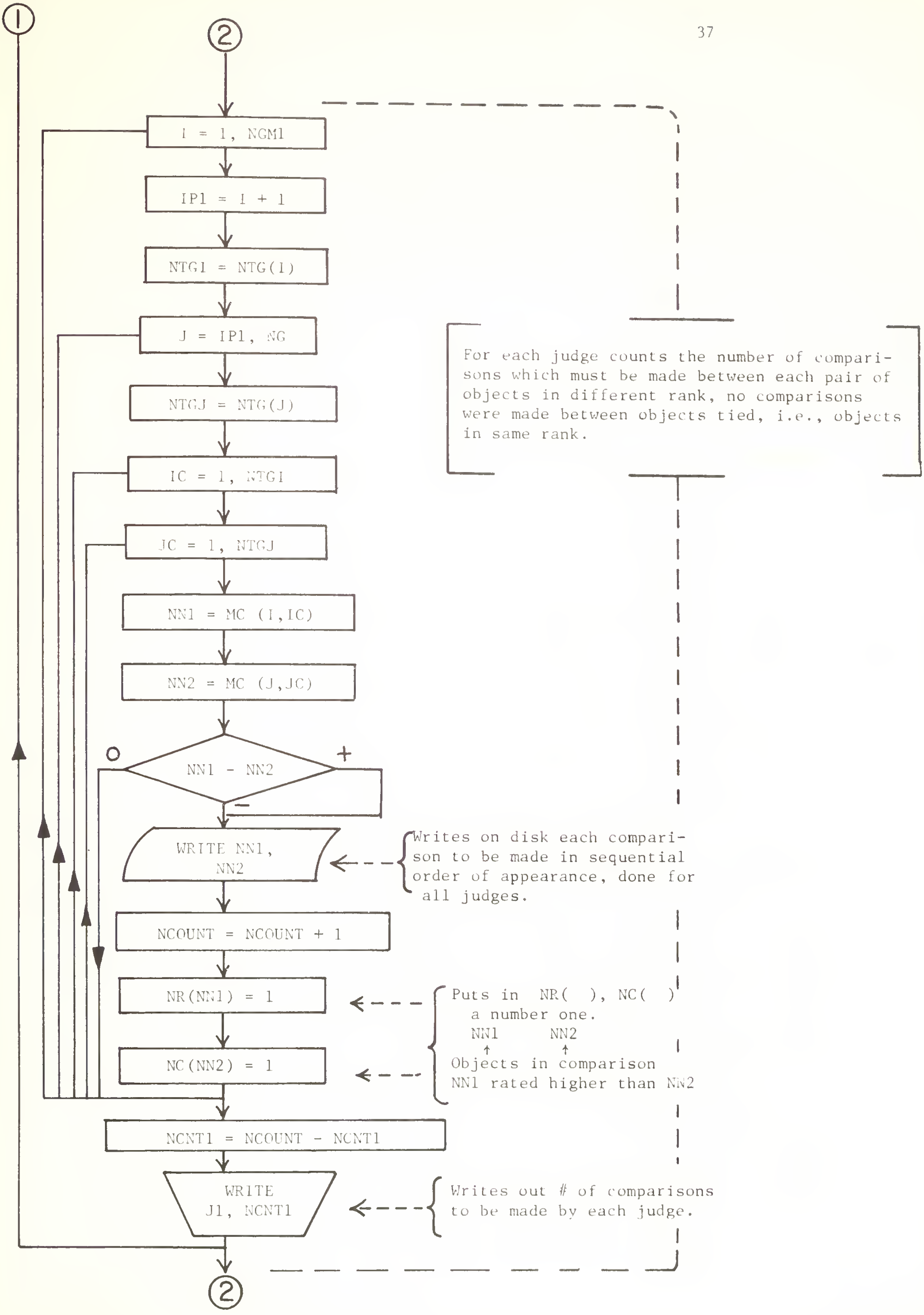
APPENDIX I

FLOW CHART OF THE FORD PROGRAM

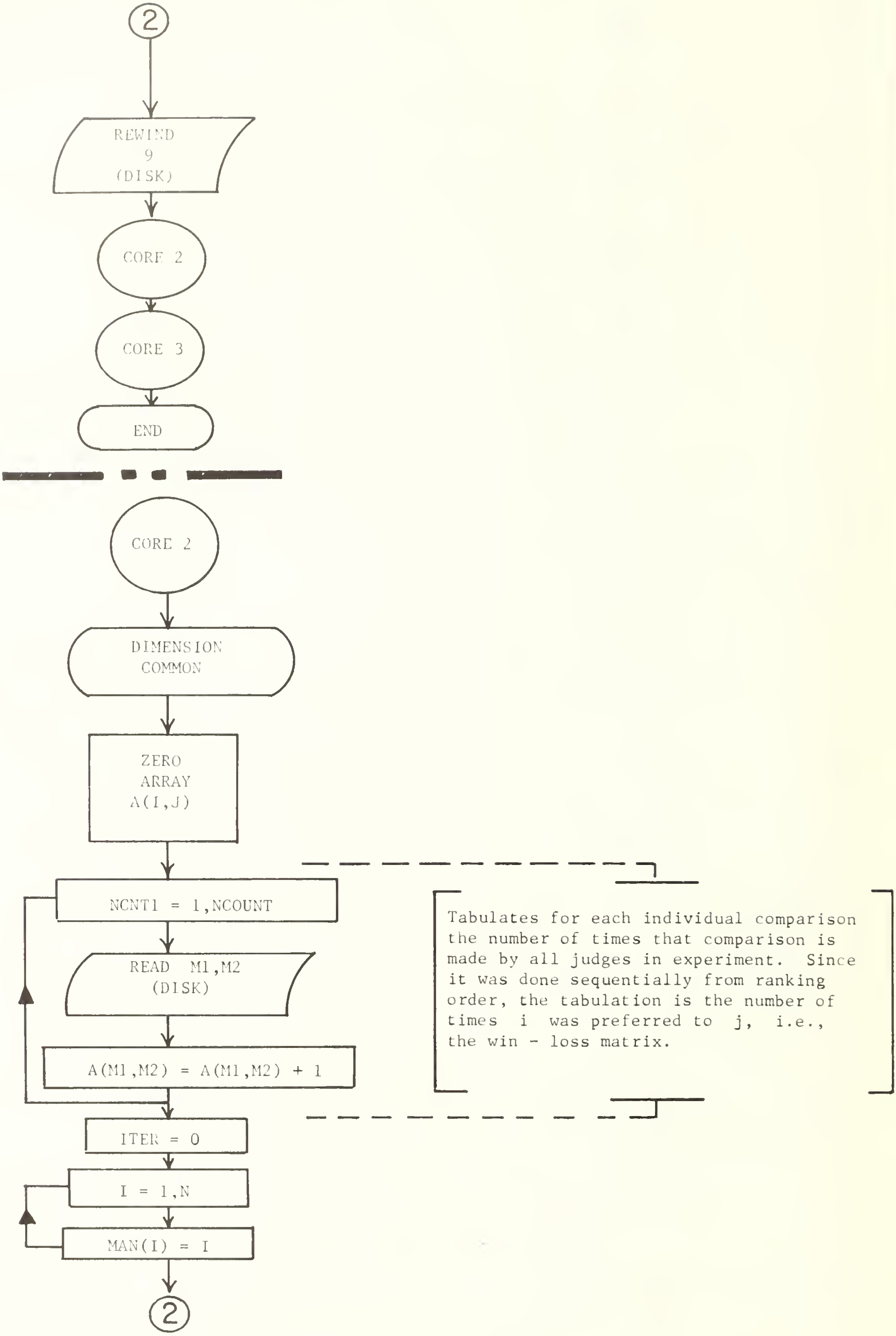

START
DIMENSION
& COMMON
READ AND
WRITE
INFO FROM
LABEL
CARD
REWIND
9
(DISK)
Zero disk for maximum usage
READ N,JJ,
EPSLON
JUPPER
Read parameter values
N = # of objects being judged
JJ = # of judges
EPSLON = convergence criterion
JUPPER = max # of iterations w/out convergence.
ZERO
NR(I) AND
NC(I)
NCOUNT = 0
Sets # of comparisons counter to zero
J1 = 1,JJ
NCNT1 = NCOUNT
READ
NG
# of ranking categories for judge being considered
2
1



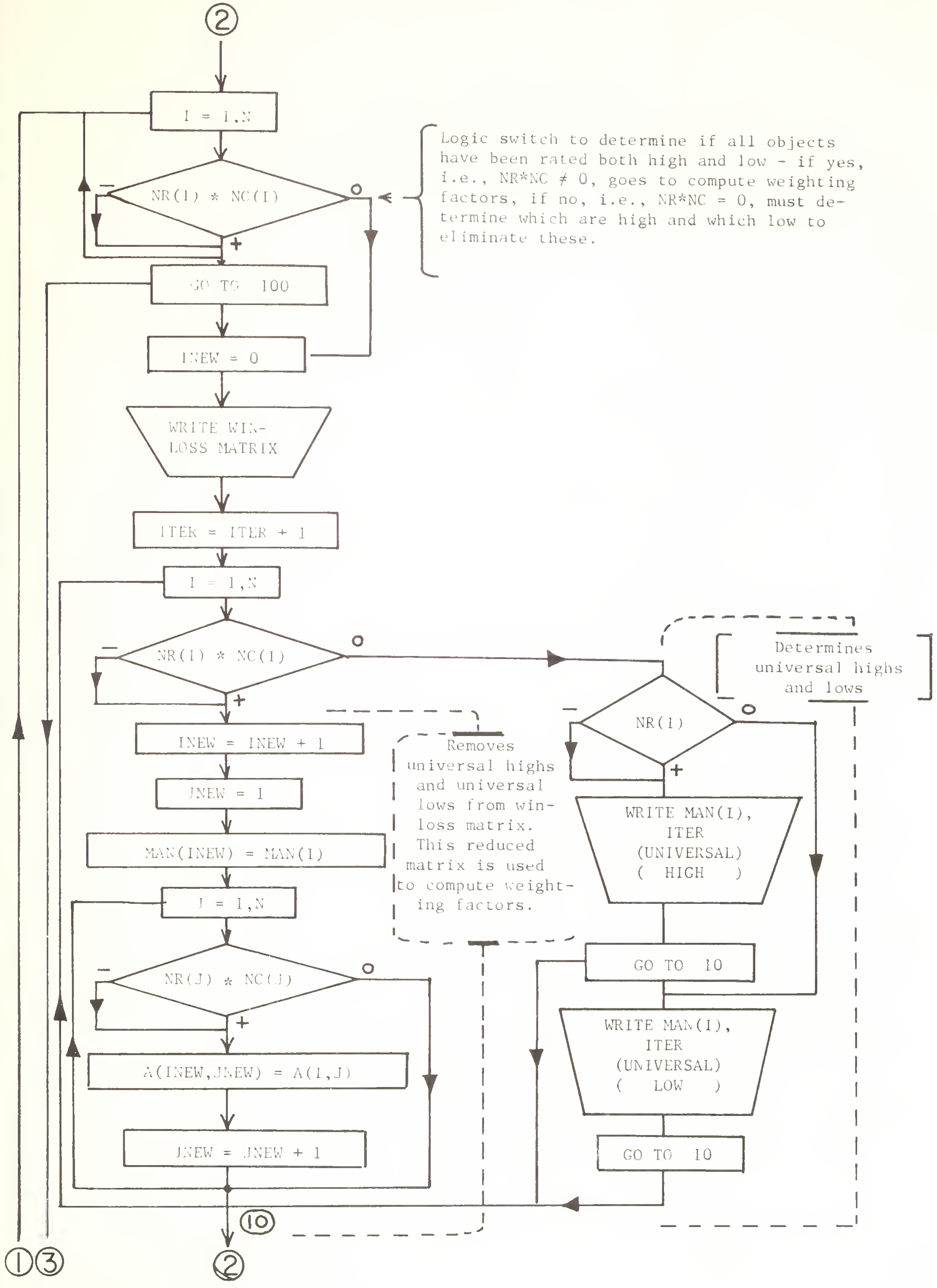
2
NG1 = 1,NG
READ NTG1, MC(NG1,K) K = 1, NTG1
Reads data cards
NTG1 = # of objects in rank
MC(,) = original ID # of objects being considered
I = 1, NTG1
LEN.EQ.0
Y
N
For each judge in turn assigns a sequential ID # to original ID # of objects being judged in order of appearance until all objects are accounted for. No duplication of assignments are made.
J = 1, LEN
MC(NG1,I).EQ.LIST(J)
Y
N
LEN = LEN + 1
LIST(LEN) = MC(NG1,I)
Puts original ID # of objects into list ( ) in order of first appearance.
J = LEN
MC(NG1,I) = J
Replaces original ID # with sequential ID #.
NTG(NG1) = NTG1
Puts into NTG( ), the # of objects in each rank category.
NGM1 = NG - 1
1
2

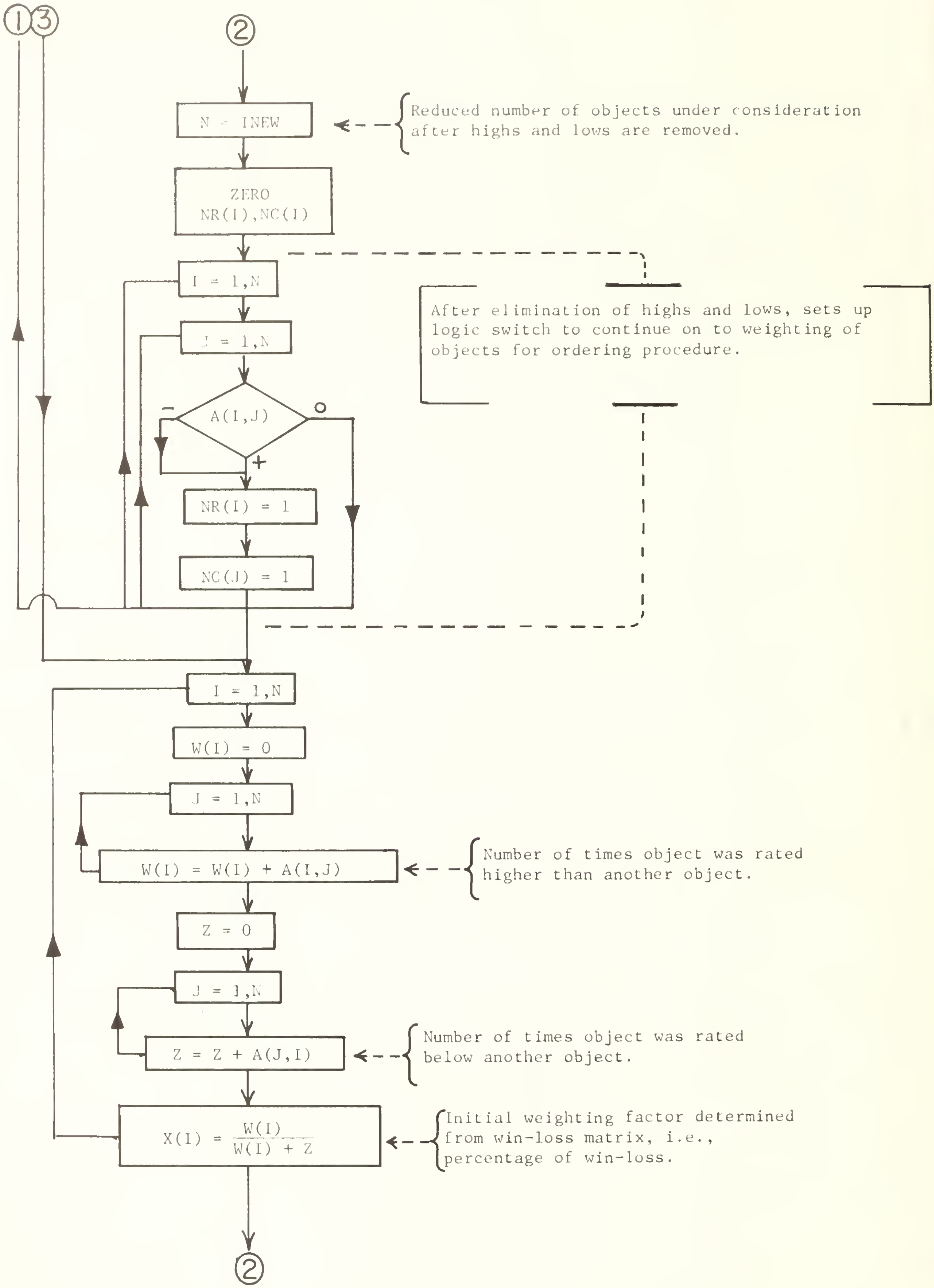


1
2
I = 1, NGM1
IP1 = I + 1
NTG1 = NTG(I)
J = IP1, NG
NTGJ = NTG(J)
IC = 1, NTGI
JC = 1, NTGJ
NN1 = MC (I,IC)
NN2 = MC (J,JC)
NN1 - NN2
O
+
-
WRITE NN1, NN2
NCOUNT = NCOUNT + 1
NR(NN1) = 1
NC(NN2) = 1
NCNT1 = NCOUNT - NCNT1
WRITE J1, NCNT1
2
For each judge counts the number of comparisons which must be made between each pair of objects in different rank, no comparisons were made between objects tied, i.e., objects in same rank.
Writes on disk each comparison to be made in sequential order of appearance, done for all judges.
Puts in NR( ), NC( ) a number one.
NN1 NN2
Objects in comparison NN1 rated higher than NN2
Writes out # of comparisons to be made by each judge.

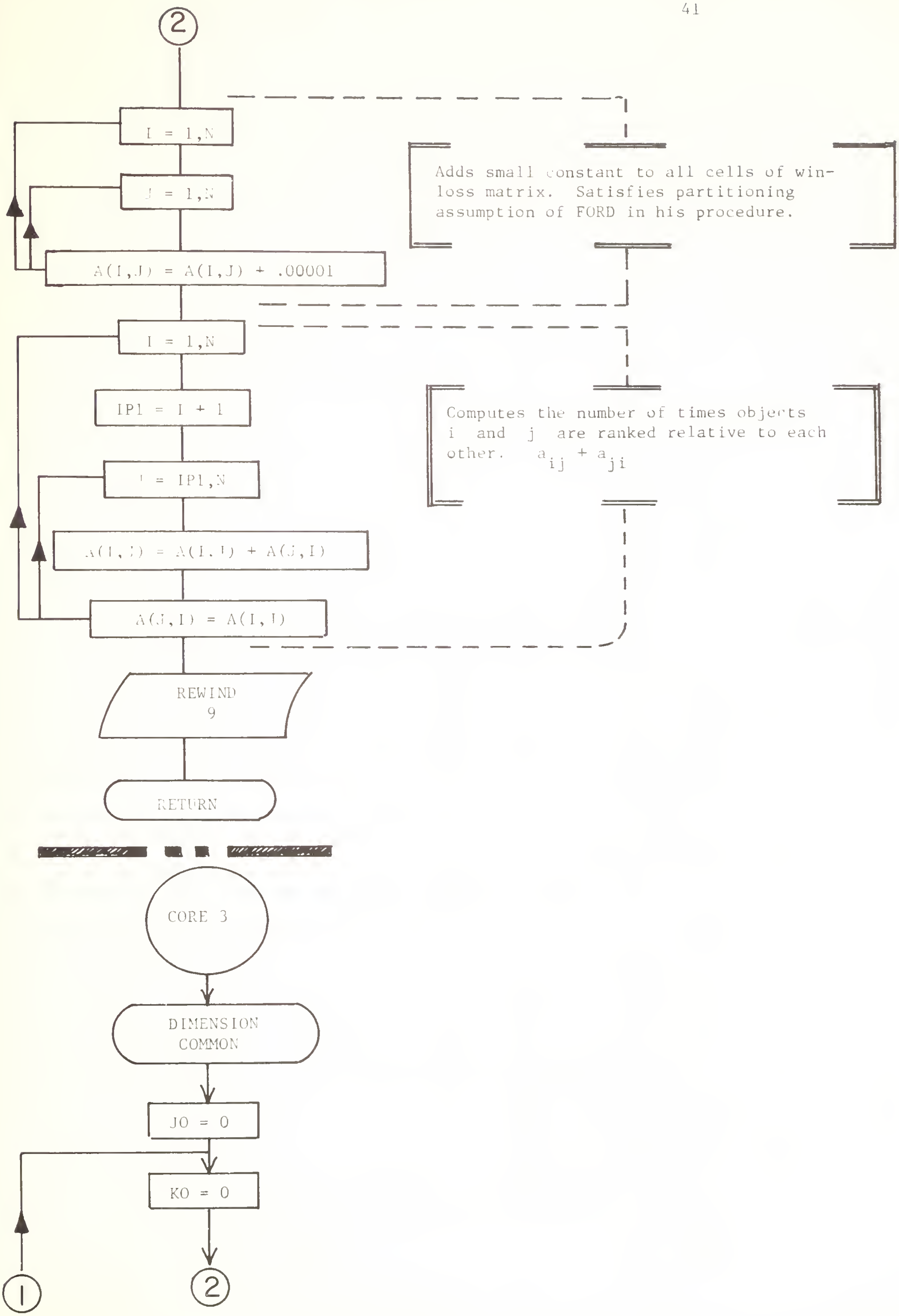


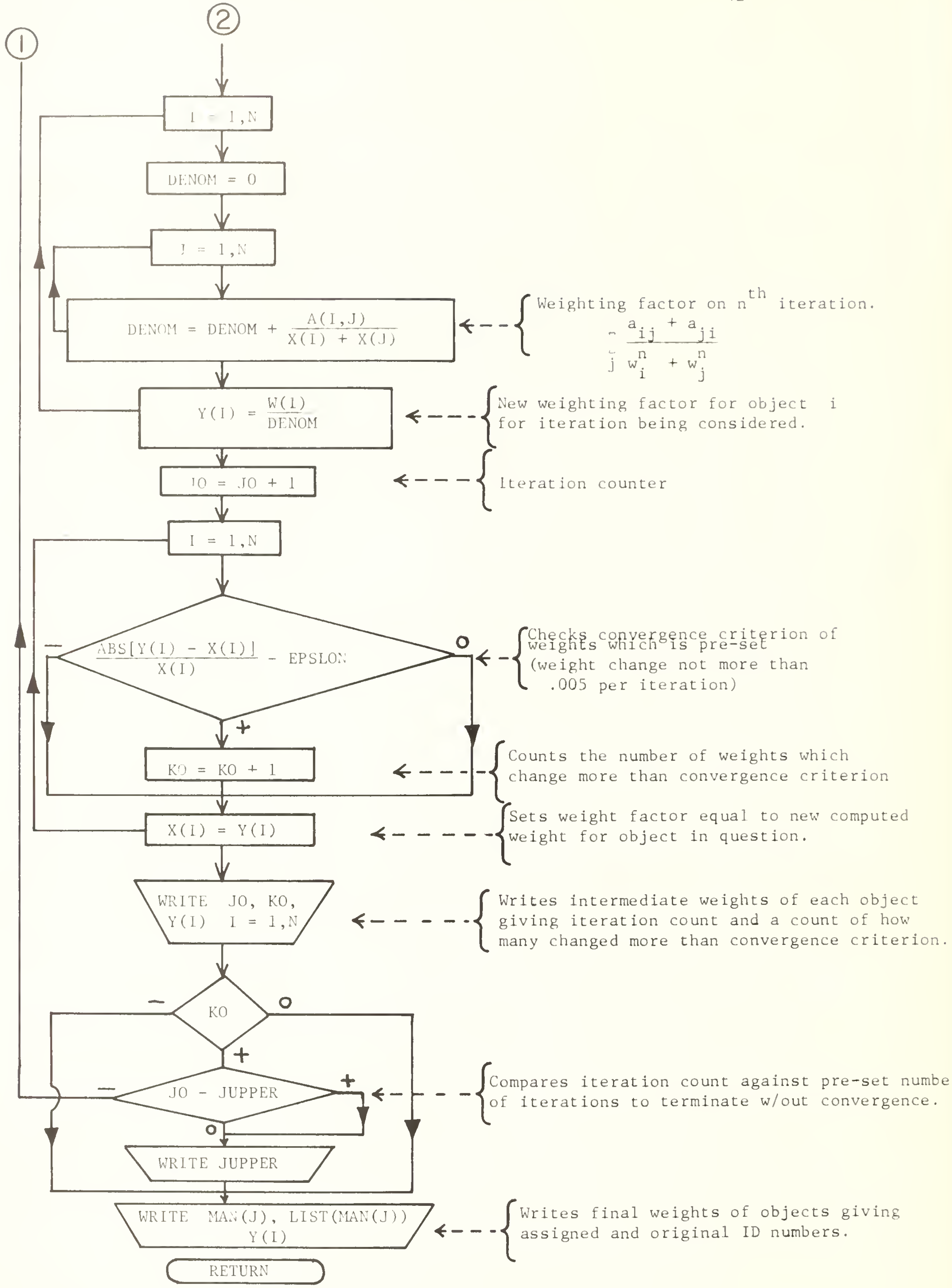
2
REWIND
9
(DISK)
CORE 2
CORE 3
END
CORE 2
DIMENSION
COMMON
ZERO
ARRAY
A(I,J)
NCNT1 = 1,NCOUNT
READ M1,M2
(DISK)
A(M1,M2) = A(M1,M2) + 1
Tabulates for each individual comparison the number of times that comparison is made by all judges in experiment. Since it was done sequentially from ranking order, the tabulation is the number of times i was preferred to j, i.e., the win - loss matrix.
ITER = 0
I = 1,N
MAN(I) = I
2


②
I = 1,N
NR(I) * NC(I)
−
+
0
Logic switch to determine if all objects have been rated both high and low - if yes, i.e., NR*NC ≠ 0, goes to compute weighting factors, if no, i.e., NR*NC = 0, must determine which are high and which low to eliminate these.
GO TO 100
INEW = 0
WRITE WIN-LOSS MATRIX
ITER = ITER + 1
I = 1,N
NR(I) * NC(I)
INEW = INEW + 1
JNEW = 1
MAN(INEW) = MAN(I)
J = 1,N
NR(J) * NC(J)
A(INEW,JNEW) = A(I,J)
JNEW = JNEW + 1
Removes universal highs and universal lows from win-loss matrix. This reduced matrix is used to compute weighting factors.
Determines universal highs and lows
NR(I)
WRITE MAN(I), ITER (UNIVERSAL) (HIGH)
GO TO 10
WRITE MAN(I), ITER (UNIVERSAL) (LOW)
GO TO 10
⑩
①③
②



1
3
2
N = INEW
Reduced number of objects under consideration after highs and lows are removed.
ZERO NR(I),NC(I)
I = 1,N
J = 1,N
After elimination of highs and lows, sets up logic switch to continue on to weighting of objects for ordering procedure.
A(I,J)
-
0
+
NR(I) = 1
NC(J) = 1
I = 1,N
W(I) = 0
J = 1,N
W(I) = W(I) + A(I,J)
Number of times object was rated higher than another object.
Z = 0
J = 1,N
Z = Z + A(J,I)
Number of times object was rated below another object.
X(I) = W(I) / W(I) + Z
Initial weighting factor determined from win-loss matrix, i.e., percentage of win-loss.
2


2
I = 1,N
J = 1,N
A(I,J) = A(I,J) + .00001
Adds small constant to all cells of win-loss matrix. Satisfies partitioning assumption of FORD in his procedure.
I = 1,N
IP1 = I + 1
J = IP1,N
A(I,J) = A(I,J) + A(J,I)
A(J,I) = A(I,J)
Computes the number of times objects i and j are ranked relative to each other. $a_{ij} + a_{ji}$
REWIND 9
RETURN
CORE 3
DIMENSION COMMON
JO = 0
KO = 0
1
2


②
①
I = 1,N
DENOM = 0
J = 1,N
DENOM = DENOM + A(I,J) / X(I) + X(J)
Weighting factor on n th iteration.
Σj (aij + aji) / (wi n + wj n)
Y(I) = W(I) / DENOM
New weighting factor for object i for iteration being considered.
JO = JO + 1
Iteration counter
I = 1,N
ABS[Y(I) - X(I)] / X(I) - EPSLON
Checks convergence criterion of weights which is pre-set (weight change not more than .005 per iteration)
KO = KO + 1
Counts the number of weights which change more than convergence criterion
X(I) = Y(I)
Sets weight factor equal to new computed weight for object in question.
WRITE JO, KO, Y(I) I = 1,N
Writes intermediate weights of each object giving iteration count and a count of how many changed more than convergence criterion.
KO
JO - JUPPER
Compares iteration count against pre-set number of iterations to terminate w/out convergence.
WRITE JUPPER
WRITE MAN(J), LIST(MAN(J)) Y(I)
Writes final weights of objects giving assigned and original ID numbers.
RETURN

APPENDIX II

DATA ASSEMBLY FOR INPUT TO THE FORD PROGRAM

DATA CARD SET UP

A. LABEL CARD - Type "1" in col. one, then any 71H. (This will be out by machine).

B. PARAMETER CARD - All numbers right adjusted. Omit all leading zeros.

Col. 1-6 - Total # of objects being compared by all judges ≤ 130

Col. 7-12 - # of judges ≤ 130

Col. 13-18 - Convergence criterion (.005 presently used)

Col. 19-24 - Max # of iterations

C. JUDGE CARD - Right adjusted. Omit leading zeros.

Col. 1-6 - # of ranks used by judge ≤ 130

D. DATA CARD - Right adjusted. Use leading zeros.

Col. 1-3 - # of objects placed in this rank by judge.

Col. 4-6 - ID # of object (original ID #)

7-9 - "

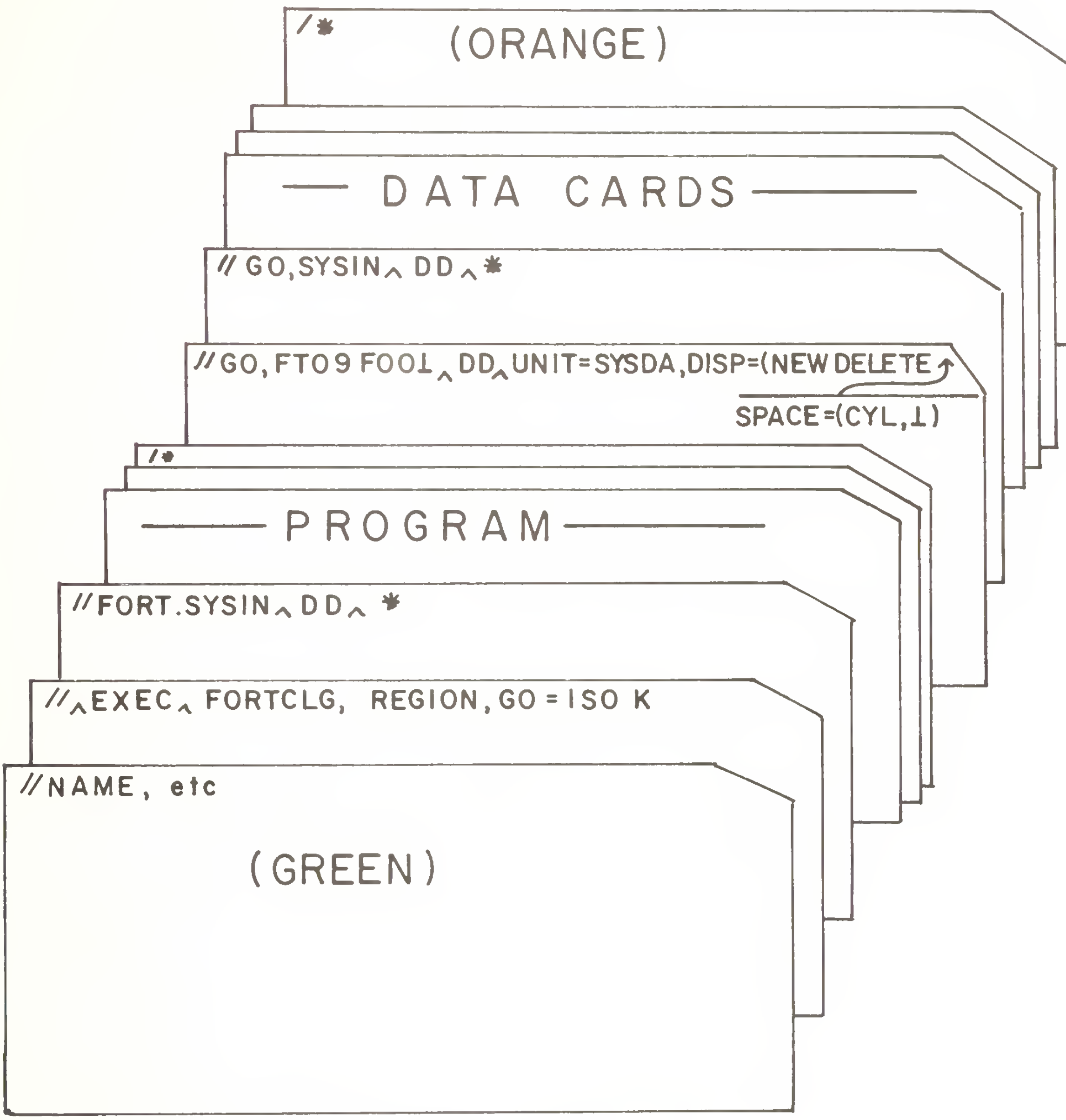
. .
. .
. .

70-72 "

Continue with as many cards as necessary to fill out rank. Subsequent cards begin ID # Col. 1-3.

Repeat C. and D. for each judge.

CARD ASSEMBLY



APPENDIX III

THE FORD PROGRAM

```
C        A COMPUTER RANKING PROCEDURE FOR PARTIALLY ORDERED          C
C        DATA. COMPARISONS MADE IN THREE CORES.                      C
C             INPUT VARIABLE NAMES-                                  C
C              N= # OF OBJECTS BEING RANKED    < OR = TO 130         C
C              JJ= # OF JUDGES    < OR = TO 130                      C
C              EPSLON= CONVERGENCE CRITERION                         C
C              JUPPER= MAX # OF ITERATION> DESIRED W/OUT             C
C                      CONVERGENCE   .                               C
C        THE DATA DECK IS ARRANGED IN THE FOLLOWING MANNER-          C
C           1.LABEL CARD- PUNCH 1 IN COL 1, FOLLOWED BY 71H          C
C           2.PARAMETER CARD- FORMAT(2I6,F6.3,I6)                    C
C             COL 1-6 = # OF OBJECTS (N)                             C
C             COL 7-12 = # OF JUDGES (JJ)                            C
C             COL 13-18 = CONVERGENCE CRITERION (EPSLON)             C
C             COL 19-24 = MAX # OF ITERATIONS (JUPPER)               C
C           3.JUDGE CARD- FORMAT(I6)                                 C
C             COL 1-6 = # OF RANKS USED BY THIS JUDGE (NG)           C
C           4.DATA CARDS- FORMAT(24I3)                               C
C             COL 1-3 = # OF OBJECTS IN THIS RANK                    C
C             COL 4-72 = ID NUMBER OF OBJECTS IN THIS RANK,          C
C          CONTINUE ON WITH AS MANY CARDS AS NECESSARY TO            C
C          COMPLETE A RANK.                                          C
C           5.THERE MUST BE (JJ) BLOCKS OF (3) AND (4).              C
C                                                                    C
      INTEGER*2 LIST(130)
      DIMENSION NR(130),NC(130),NTG(130),MC(130,130)
      COMMON LIST
      COMMON NG,NGM1,IP1,NTGI,NTGJ,NTG,EPSLON,NR,NC,N,NCOUNT,JUPPER,MC
      INTEGER OUTPUT
      LEN=0
      INPUT=5
      IDISK=9
      OUTPUT=6
      READ(INPUT,1000)
 1000 FORMAT (72H
     X                    )
      WRITE(OUTPUT,1000)
      REWIND 9
      READ(INPUT,1) N,JJ,EPSLON,JUPPER
    1 FORMAT (2I6,F6.3,I6)
      DO 2 I=1,N
      NR(I)=0
    2 NC(I)=0
      NCOUNT=0
      DO 8 J1=1,JJ
      NCNT1=NCOUNT
      READ(INPUT,3) NG
    3 FORMAT (I6)
```

```
C                                                                          C
C                                                                          C
C         A SEQUENTIAL ID NUMBER IS ASSIGNED TO THE ORIGINAL                C
C         ID NUMBER OF THE OBJECTS BEING JUDGED IN ORDER OF                 C
C         THEIR APPEARANCE IN THE DATA, UNTIL ALL OBJECTS                   C
C         BEING JUDGED ARE ACCOUNTED FOR. NO DUPLICATION OF                 C
C         ASSIGNMENTS IS MADE.                                              C
C                                                                          C
      DO 5 NG1=1,NG
      READ(INPUT,4) NTG1,(MC(NG1,K),K=1,NTG1)
    4 FORMAT (24I3,8X)
      DO 10 I=1,NTG1
      IF(LEN.EQ.0) GO TO 25
      DO 20 J=1,LEN
      IF(MC(NG1,I).EQ.LIST(J)) GO TO 30
   20 CONTINUE
   25 LEN=LEN+1
      LIST(LEN)=MC(NG1,I)
      J=LEN
   30 MC(NG1,I)=J
   10 CONTINUE
    5 NTG(NG1)=NTG1
      NGM1=NG-1
C                                                                          C
C         THE NUMBER OF COMPARISONS WHICH MUST BE MADE BY                   C
C         EACH JUDGE IS COMPUTED. NO COMPARISONS ARE MADE                   C
C        BETWEEN OBJECTS IN THE SAME RANK.                                  C
C                                                                          C
      DO 7 I=1,NGM1
      IP1=I+1
      NTGI=NTG(I)
      DO 7 J=IP1,NG
      NTGJ=NTG(J)
      DO 7 IC=1,NTGI
      DO 7 JC=1,NTGJ
      NN1=MC(I,IC)
      NN2=MC(J,JC)
      IF (NN1-NN2) 6,7,6
    6 WRITE(IDISK) NN1,NN2
      NCOUNT=NCOUNT+1
      NR(NN1)=1
      NC(NN2)=1
    7 CONTINUE
      NCNT1=NCOUNT-NCNT1
    8 WRITE(OUTPUT,9) J1,NCNT1
    9 FORMAT ( 6H0JUDGE I5, 10(1H ),I7,12H COMPARISONS )
      END FILE 9
      REWIND 9
```

```
      WRITE(OUTPUT,500)
  500 FORMAT(1H0,'ID NUMBER MAP')
      WRITE (OUTPUT,501)
  501 FORMAT(15(1H ),13HASSIGNED ID #,15(1H ),13HORIGINAL ID #/)
      WRITE(OUTPUT,502) (I,LIST(I),I=1,N)
  502 FORMAT(T20,I4,T49,I4)
      CALL CORE2
      CALL CORE3
      STOP
      END
C                                                                        C
C          RANKING PROCEDURE -- CORE2                                    C


      SUBROUTINE CORE2
      INTEGER*2 LIST(130)
      DIMENSION NR(130),NC(130),A(130,130),W(130),X(130),MAN(130)
      COMMON LIST
      COMMON M1,M2,INEW,JNEW,IP1,W,EPSLON,NR,NC,N,NCOUNT,JUPPER,A,X,MAN
      INTEGER OUTPUT
      INPUT=5
      OUTPUT=6
      IDISK=9
      DO 1 I=1,N
      DO1 J=1,N
    1 A(I,J)=0
C                                                                        C
C          TABULATES FOR EACH COMPARISON THE NUMBER OF TIMES             C
C          THAT COMPARISON IS MADE BY ALL JUDGES IN THE                  C
C          EXPERIMENT. THIS TABULATION IS DONE FROM RANKING              C
C          ORDER AND IS THE NUMBER OF TIMES OBJECT I WAS                 C
C          PREFERRED TO OBJECT J,IE,THE WIN-LOSS MATRIX.                 C
C                                                                        C
      DO 2 NCNT1=1,NCOUNT
      READ(IDISK) M1,M2
    2 A(M1,M2)=A(M1,M2)+1
      ITER=0
      DO 34 I=1,N
   34 MAN(I)=I
C                                                                        C
C          OBJECTS WHICH ARE RATED UNIVERSALLY HIGH OR                   C
C          UNIVERSALLY LOW ARE REMOVED FROM THE COMPUTATIONS.            C
C          THE APPROPRIATE ROWS AND COLUMNS OF THE WIN-LOSS              C
C          MATRIX ARE ALSO REMOVED.                                      C
C                                                                        C
   33 DO 3 I=1,N
      IF (NR(I)*NC(I)) 3,4,3
    3 CONTINUE
```

```
      GO TO 100
    4 INEW=0
      WRITE(OUTPUT,300)
  300 FORMAT(1HO,'WIN-LOSS MATRIX ORDERED BY ASSIGNED ID NUMBER')
      DO 70 I=1,N
   70 WRITE(OUTPUT,36) (A(I,J),J=1,N)
   36 FORMAT(30F4.0)
      ITER=ITER+1
      DO 10 I=1,N
      IF(NR(I)*NC(I)) 9,18,9
    9 INEW=INEW+1
      JNEW=1
      MAN(INEW)=MAN(I)
      DO 1000 J=1,N
      IF (NR(J)*NC(J)) 11,1000,11
   11 A(INEW,JNEW)=A(I,J)
      JNEW=JNEW+1
 1000 CONTINUE
   10 CONTINUE
      N=INEW
      DO 31 I=1,N
      NR(I)=0
   31 NC(I)=C
      DO 30 I=1,N
      DO 30 J=1,N
      IF (A(I,J)) 32,30,32
   32 NR(I)=1
      NC(J)=1
   30 CONTINUE
      GO TO 33
C                                                                     C
C                                                                     C
C          COMPUTES THE INITIAL WEIGHTING FACTORS OF THE             C
C          OBJECTS FROM THE WIN-LOSS MATRIX. THESE WEIGHTING         C
C          FACTORS ARE THE PROBABILITY THAT OBJECT I IS
C          PREFERRED TO OBJECT J.                                     C
C                                                                     C
  100 DO 16 I=1,N
      W(I)=0
      DO 13 J=1,N
   13 W(I)=W(I)+A(I,J)
      Z=0
      DO 15 J=1,N
   15 Z=Z+A(J,I)
   16 X(I)=W(I)/(W(I)+Z)
      DO 200 I=1,N
      DO 200 J=1,N
  200 A(I,J)=A(I,J)+ .00001
```

```
C                                                                      C
C         COMPUTES THE NUMBER OF TIMES OBJECT I AND OBJECT J           C
C         ARE RANKED RELATIVE TO EACH OTHER.                           C
C                                                                      C
      DC 14 I=1,N
      IP1=I+1
      DO 14 J=IP1,N
      A(I,J)=A(I,J)+A(J,I)
   14 A(J,I)=A(I,J)
      REWIND 9
      RETURN
   18 IF (NR(I)) 8,5,8
    8 WRITE(OUTPUT,12) MAN(I),ITER
   12 FCRMAT(1HO,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED HI
     XGH,DELETED,I3, 3H RY )
      GC TO 10
    5 WRITE(OUTPUT,7) MAN(I),ITER
    7 FORMAT(1HO,22HSUBJECT ASSIGNED ID #,I3,35H IS UNIVERSALLY RATED LO
     XW ,DELETED,I3, 3H RY )
      GO TO 10
      END
C                                                                      C
C         RANKING PROCEDURE -- CORE3                                   C
C                                                                      C


      SUBROUTINE CORE3
      INTEGER*2 LIST(130)
      DIMENSION NR(130),NC(130),A(130,130),W(130),X(130),MAN(130),
     1 Y(130),Q(130)
      CCMMON LIST
      CCMMON JO,KC,DENOM,XA,XQ,W,EPSLCN,NR,NC,N,NCOUNT,JUPPER,A,X,MAN
      INTEGER OUTPUT
      INPUT=5
      OUTPUT=6
      IDISK=9
      JO=0
    8 KC=0
C                                                                      C
C         COMPUTES NEW WEIGHTING FACTOR FOR OBJECT I FOR               C
C         ITERATICN BEING CONSIDERED.                                  C
C                                                                      C
      DO 3 I=1,N
      DENOM=C
      CO 4 J=1,N
    4 DENOM=DENOM+A(I,J)/(X(I)+X(J))
    3 Y(I)=W(I)/DENOM
      JO=JO+1
```

```
C                                                                  C
C         CONVERGENCE CRITERION CHECH.                             C
C                                                                  C
    5 DC 9 I=1,N
      IF(ABS(Y(I)-X(I))/X(I)-EPSLON) 9,9,10
   10 KO=KO+1
    9 X(I)=Y(I)
      WRITE(OUTPUT,12) JO,KO,(Y(I),I=1,N)
   12 FCRMAT (1HO,I5,I10/(6F18.6))
      IF (KO) 17,17,13
C                                                                  C
C         ITERATICN COUNT CHECK.                                   C
C                                                                  C
   13 IF (JO-JUPPER) 8,14,14
   14 WRITE(OUTPUT,15) JUPPER
   15 FORMAT(1HO,I5,46H ITERATIONS, NO CONVERGENCE. DATA SET DELETED. )
   17 WRITE(OUTPUT,18) (MAN(I),LIST(MAN(I)),Y(I),I=1,N)
   18 FCRMAT (14H1FINAL WEIGHTS//15(1H ),
     X13HASSIGNED ID #, 3(1H ),13HORIGINAL ID #,12(1H ),6HWEIGHT//(1HO,
     X2I19,F24.6))
      RETURN
      END
```

APPENDIX IV

PROJECT DESCRIPTIONS FOR TRIAL APPLICATION TESTING

PROJECT DESCRIPTIONS FOR TRIAL APPLICATION TESTING

1. TITLE: Improved Enlisted Personnel Distribution and Management.

DESCRIPTION: A computer assisted distribution and assignment (CADA) system is being designed to help improve the utilization of enlisted manpower. Preliminary model currently is being implemented in the Pacific Fleet. Prototype model is now under development for application in BUPERS in support of centralized management of enlisted ratings. Related research results include development of computer and mathematically based procedures for (1) the equitable allocation of personnel resources, (2) the optimal match of man and billet, (3) the identification of billet vacancies in order of priority, (4) the projection of the number of distributable assets, and (5) the feedback of information on the results of distribution management actions.

2. TITLE: Ship Manning Requirements Techniques

DESCRIPTION: The increasing sophistification and complexity of naval ships, systems, and equipments in the face of project volunteer and a smaller Navy requires the development of methods which will improve the accuracy of manpower requirements forecasting and manpower utilization.

A technique for defining and documenting manpower requirements for ships based on the application of selected work study techniques to basic manning criteria in each of the separate work areas aboard ship has been developed. It permits the production of a document which displays in detail the rationale for manning by ship classes based on equipment and required operational capabilities to meet mission assignment.

3. TITLE: Evaluation of Standards for Navy Reenlistment.
DESCRIPTION: This research was generated out of concern over the quality of reenlistees. Unsatisfactory performance was costing the military services enormous amounts of money in such things as reenlistment bonuses and pay and allowances for reenlistees from whom commensurate service was not realized. Court and confinement costs of reenlistees were cited. It was suspected that personnel of inferior quality were being allowed to reenlist, including some with unsatisfactory first term records.

In an attempt to identify unsatisfactory individuals prior to reenlistment, comparisons were made between unsatisfactory and satisfactory reenlistees on information available at the time of the reenlistment decision. The project also provided information on the effect on manning which would result if reenlistment standards were made more stringent.

4. TITLE: Development of Navy Military Personnel Costing Techniques for Use in Determining Cost Implications Associated with Changes in Reenlistment Rates.

DESCRIPTION: Thousands of skilled technicians are required to operate and maintain the complex systems and equipment now in the Fleet. The Navy constantly experiences difficulty in retaining these technicians because of competition for them from other sectors of the economy.

To alleviate this problem, several technician-oriented procurement programs and career incentive programs are employed. To facilitate evaluation of these programs, a methodology for determining the relative cost benefits associated with retention of personnel has been developed.

5. TITLE: Design of an Optimum Personnel Force Structure.

DESCRIPTION: An optimum force structure containing appropriately qualified personnel in sufficient numbers at least cost cannot now be certified. This project is concerned with the development of improved techniques to analyze and balance the relationship between personnel requirements and the composition of the existing force structure.

6. TITLE: Interest Measurement in Officer Selection.

DESCRIPTION: Each year several thousand young men apply for officer training programs at the Naval Academy and NROTC units at various colleges. High attrition rates

are experienced in both training and active duty. To reduce the cost of losing substantial proportions of these men, it is imperative that those applicants having the greatest career potential be identified in the selection process. Several years of research on vocational interest tests and biographical questionnaires have made it possible to identify those applicants most likely to successfully complete officer training and remain in the Navy after completing their minimum requirements.

7. TITLE: Evaluation Survey of the Effectiveness of Submarine Sonar Operator Training.

DESCRIPTION: A comprehensive survey was accomplished of the proficiency, training, and utilization of submarine sonar technicians and sonar watchstanders. The survey provided up-to-date information concerning the efficiency of training procedures. Such information is necessary on a periodic basis to insure appropriate alignment of the training to fleet requirements in order to prevent serious impairment of operational fleet submarine ASW efficiency. Data gathering instruments included interview forms, self ratings, supervisor ratings, knowledge tests, and performance tests.

8. TITLE: Marginal Personnel/Minority Group Testing.

DESCRIPTION: Present test batteries used in both military and civilian settings have been criticized for alleged inequities when used with groups defined on the basis of race or ethnic affiliation. Public policy as well as

efficient manpower utilization requires that all personnel be afforded equality of opportunity in assignment and that those abilities being measured bear relevance to skills required on-the-job.

9. TITLE: Personnel Cost Research for Early Man/Machine Design Trade-Offs.

DESCRIPTION: The critical element of personnel cost has not been systematically considered when making system design and development decisions early in the system development cycle. No tools exist to enable the cost-effectiveness of such decisions to be measured. For this reason, research was undertaken to develop a personnel cost model for use in personnel and man-equipment trade off decisions. A basis model was accomplished which allowed the identification of all pertinent cost items and the accumulation of cost elements in an unequivocal manner.

10. TITLE: LOFARGRAM Analysis Procedures.

DESCRIPTION: The airborn JEZEBEL system has shown great potential as a means of detecting and classifying underwater contacts; however, its usefulness has been continually hampered by the lack of adequately trained operators. One of the main reasons for operator deficiencies is that training programs have been seriously hampered by the lack of a standardized, systemic procedure for analyzing the information displayed on the gram which is the main display component of the system.

In order to correct this situation, a systematic LOFARGRAM procedure was developed.

DISTRIBUTION LIST

	No. Copies
Personnel Research Division (Pers-A3) Bureau of Naval Personnel Washington, D. C. 20390	50
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12
Dean of Research Administration Code 023 Naval Postgraduate School Monterey, California 93940	2
Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
Professor James K. Arima Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	30
Professor J. R. Borsting Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
Library Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	2
Psychology Research Unit Australian Military Forces Albert Parks Barracks Melbourne, Australia	1
Dr. J. D. Nelms, Director Army Personnel Research Establishment c/o Royal Aircraft Establishment Farnborough, Hants. United Kingdom	1

	No. Copies
Burton V. Dean Chairman, Operations Research Department School of Management Case Western Reserve University University Circle Cleveland, Ohio 44106	1
Philip R. Shelton North Devon College Barnstaple Devon, ENGLAND	1
Dr. Curtis J. Tompkins Graduate School of Business Administration University of Virginia Box 3607 Charlottesville, Virginia 22903	1
Professor Chester C. Kisiel Hydrology and Water Resources Office University of Arizona Tucson, Arizona 85721	1
W. Herroelen Institute of Applied Economics Leuven University Dekenstraat 2 B - 3000 Leuven (BELGIUM)	1
Mr. Pierre Villani Documentation Centre Data Processing Centre Ville De Montreal City of Montreal Edifice Bonsecours 330 Est. Rue St.-Paul Montreal, CANADA	1
D. J. Luaghhunn Department of Administrative Sciences Southern Illinois University at Carbondale Carbondale, Illinois 62901	1
John C. Dyer Group Leader AVCO Government Products Group 201 Lowell Street Wilmington, Mass. 01887	1

	No. Copies
J. A. Palermo 1449 Greenmont Ct. Reston, Virginia 22070	1
Dr. Roy B. Larson Manager, Strategic Planning Scott Paper Company Scott Plaza Philadelphia, Pennsylvania 19113	1
Dr. Burt Nanus, Director Graduate School of Business Administration University of Southern California University Park Los Angeles, California 90007	1
Professor Kenneth L. Kraemer Graduate School of Administration University of California Irvine, California 92664	1
Y. Y. Yen Department of Quantitative Methods Graduate School of Business The University of Santa Clara Santa Clara, California 95053	1
Professor Paul E. Green Department of Marketing and International Business Wharton School of Finance and Commerce University of Pennsylvania Philadelphia, Pennsylvania 19104	1
R. Edward Overstreet Faculty of Administrative Studies Room N.711, Ross Building York University 4700 Keele Street Downsview, Ontario, CANADA	1
James M. Howell Allstate The Judson Branch Research Center 321 Middlefield Road Menlo Park, California 94025	1

	No. Copies
Professor James S. Dyer Graduate School of Business Administration University of California Los Angeles, California 90024	1
Mr. Meredith A. Gonyea The Ohio State University College of Administrative Science 1775 South College Road Columbus, Ohio 43210	1
Professor Bertrand Horwitz Syracuse University School of Management 116 College Place Syracuse, New York 13210	1
Professor Fred R. McFadden School of Business University of Colorado Colorado Springs Center Cragmor Road Colorado Springs, Colorado 80907	1
Professor Robert D. Dahle Agricultural Extension Service North Carolina State University School of Agriculture and Life Sciences Box 5576 Raleigh, North Carolina 27607	1
B. E. Gillett Computer Science Department University of Missouri-Rolla Rolla, Missouri 65401	1
Arthur Yates Glidden-Durkee Executive Offices 900 Union Commerce Building Cleveland, Ohio 44115	1
Professor Paul Gray School of Industrial and Systems Engineering Georgia Institute of Technology Atlanta, Georgia 30332	1

	No. Copies
CDR R. R. Jefferson National Cash Register Company Dayton, Ohio 45409	1
Dr. Kamal E. Said Assistant Professor College of Business Administration University of Houston Cullen Boulevard Houston, Texas 77004	1
Mr. Nelson B. Heller Staff Member Board of Police Commissioners Metropolitan Police Department St. Louis, Missouri 63100	1
Jeffrey Sidney Faculty of Commerce University of British Columbia Vancouver 8, CANADA	1
Arthur H. G. Mittermaier Ministry of the Environment 135 St. Clair Ave., West Toronto 7, Ontario, CANADA	1
J. B. Schneider Dept. of Urban Planning (JO-40) University of Washington Seattle, Washington 98195	1
J. B. Hartley Dept. of Environment Ottawa, CANADA	1
Don Grossman Johnson & Johnson, M.I.C. Raritan, New Jersey 08869	1
Doug Schweitzer Institute for Northern Studies University of Saskatchewan Saskatoon, Sask, CANADA S7N OWO	1

	No. Copies
Professor J. E. Dovley Faculty of Management Studies University of Toronto 246 Bloor St. West Toronto 181, CANADA	1
Professor Charles Pinkus Department of Operations Research The George Washington University School of Engineering and Applied Science Washington, D. C. 20006	1
R. Stanley LaVallee Procurement and Systems Acquisition Division United States General Accounting Office Washington, D. C. 20548	1
M. E. Melnyk Program Office Department of Communications 100 Metcalfe Street, Room 317 Ottawa, K1A 0C8 Ontario, CANADA	1
Dr. Alexander R. Craw Technical Analysis Division A160, 225 U. S. Department of Commerce National Bureau of Standards Washington, D. C. 20234	1
Dennis S. Skoff Engineering SDS Hough Division International Harvester Company East Synnyside Avenue Libertyville, Illinois 60048	1
Professor C. R. Klasson College of Business Administration The University of Iowa Iowa City, Iowa 52240	1
R. G. Chamberlain Jet Propulsion Laboratory California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91103	1

	No. Copies
Donald F. Phillips Science Information Program Department of Humanities Illinois Institute of Technology Chicago, Illinois 60616	1
Miss P. Sharp, Librarian ICL House Putney LONDON SW15	1
Phillip T. Drotning Director of Urban Affairs Standard Oil Company 910 South Michigan Ave. Chicago, Illinois 60680	1
K. J. Tiahrt Department of Mathematics Montana State University Bozeman, Montana 59715	1
D. A. Haith Riley-Robb Hall Cornell University Ithaca, New York 14850	1
Nancy Hales Oklahoma Legislative Council 5th Floor, Oklahoma State Capitol Oklahoma City, Oklahoma 73105	1
Paul Slovic Oregon Research Institute P. O. Box 3196 Eugene, Oregon 97403	1
Nancy S. McDonnell Research Assistant The Pennsylvania State University Institute for Research on Human Resources 407 Graduate Building University Park, Pennsylvania 16802	1

	No. Copies
David A. Lopex Dept. of Quantitative Business Analysis University of Southern California School of Business Administration University Park Los Angeles, California 90007	1
Richter Data Processing Subdivision Operations Research Deutsche Lufthansa AG 600 Frankfurt A.M. Flughafen GERMANY	1
David L. Ford The University of Wisconsin College of Engineering Dept. of Industrial Engineering 1513 University Avenue Madison, Wisconsin 53706	1
Dr. Frank M. Andrews Survey Research Center Institute for Social Research The University of Michigan Ann Arbor, Michigan	1
Professor Michael Radnor Graduate School of Management Northwestern University Evanston, Illinois 60201	1
Dr. Chia Shun Shih Department of Industrial Engineering Texas A&M University College of Engineering College Station, Texas 77843	1
Mrs. Agnes D. Livingood Director of Urban Systems Urban Studies Center Gardencourt Campus Alta Vista Road University of Louisville Louisville, Kentucky 40205	1
E. G. Dupnick, FM3 National Aeronautics and Space Administration Manned Spacecraft Center Houston, Texas 77058	1

	No. Copies
S. B. Thiebeauld Documentaliste Societe D'Informatique De Conseils Et De Recherche Operationnelle SINCRO 47, Rue Henri Heine PARIS - 16	1
Medard Z. Bryll Vice President Gateway Industries, Inc. 8825 South Greenwood Avenue Chicago, Illinois 60619	1
Commanding General USA Combat Developments Command Experimentation Command Fort Ord, California 93941	1
USA Behavioral and Systems Research Laboratory Room 239, The Commonwealth Building 1320 Wilson Boulevard Arlington, Virginia 22209	1
Commanding Officer USA Human Engineering Laboratories Aberdeen Proving Ground, Maryland 21005	1
Commanding General USA Medical Research Laboratory Fort Knox, Kentucky 40121	1
US Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235	1
Director Personnel Research Laboratory Lackland AFB, Texas 78236	1
Human Engineering Division Naval Training Device Center Orlando, Florida 32013	1
Naval Aerospace Medical Institute Pensacola, Florida 32512	1
Naval Medical Research Laboratory Submarine Base New London, Connecticut 06342	1

	No. Copies
Director Human Resources Research Organization 300 North Washington Street Alexandria, Virginia 22314	1
Training Research Distributor US Air Force Human Resources Laboratory Wright-Patterson AFB, Ohio 45433	1
US Naval Personnel Research and Development Laboratory Washington, D. C. 20390	1
US Naval Personnel and Training Research Laboratory San Diego, California 92152	1
Commander (Code 4011) Naval Weapons Center China Lake, California 93555	1
Human Performance Branch NASA Ames Research Center Moffett Field, California 94035	1
Director Human Resources Research Organization P.O. Box 5787 Presidio of Monterey Monterey, California 93940	1
Human Engineering Division Aerospace Medical Laboratories Wright-Patterson Air Force Base Attn: MRHF Dayton, Ohio 45433	1
Department of the Air Force Department of Psychology and Leadership USAF Academy, Colorado 80840	1
Director, Office of Military Psychology and Leadership West Point, New York 10996	1
Commanding Officer Naval Medical Field Research Laboratory Camp Lejeune, North Carolina 28542	1

	No. Copies
US Naval Missile Test Center Attn: Human Factors Branch Point Mugu, California 93041	1
Psychological Sciences Division (Code 450) Office of Naval Research Department of the Navy Arlington, Virginia 22217	1
US Army Enlisted Evaluation Center Fort Benjamin Harrison Indianapolis, Indiana 46249	1
Head, Human Engineering Laboratory Naval Electronics Laboratory Center San Diego, California 92152	1
Head, Aerospace Operations Psychology Branch Bureau of Medicine and Surgery Department of the Navy Washington, D. C. 20390	1
US Army Air Defense Human Research Unit Fort Bliss, Texas 79916	1
Dr. Eugene Gloye Office of Naval Research Pasadena Branch Office 1030 East Green Str-et Pasadena, California 91101	1
Motivation and Training Laboratory USA Manpower Resources Research and Development Center 1300 Wilson Boulevard Arlington, Virginia 22209	1
Professor Stefan D. Bloomfield School of Business and Technology Oregon State University Corvalis, Oregon 97331	1

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3 REPORT TITLE

Evaluation and Innovation in the Navy's Personnel Research Laboratories, II. Development, Validation, and Trial Application of a Computer Program to Facilitate Judgmental Appraisal

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report 1972

5 AUTHOR(S) *(First name, middle initial, last name)*

James K. Arima and Richard W. Mister

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
May 1972		12

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO PO-1-0001 Naval Personnel Research and Development Laboratory	NPS55AA72051A
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10 DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY Personnel Research Division (Pers-A3), Bureau of Naval Personnel

13 ABSTRACT

The previous study in this series showed that evaluation of R&D activities rests eventually on qualitative judgments. The purpose of this study was to develop, validate, and test apply a procedure for obtaining qualitative judgments economically and efficiently. The Ford procedure for scaling partially ordered sets of rankings was programmed and validated using an abstract judgmental task with an extrinsic criterion. It was given a trial application requiring the ordering on merit of current personnel research projects. Both validation and trial application results were highly satisfactory. It was concluded that the Ford procedure could be used to obtain scaled qualitative judgments in a wide variety of settings with accuracy, efficiency, and economy. Flow charts, data setup, and the complete computer program are given.

This research was supported in part by the Personnel Research Division, Bureau of Naval Personnel, through Project Order No. 1-0001, Naval Personnel Research and Development Laboratory.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Research Management						
Development Management						
Laboratory Effectiveness						
In-House Laboratories						
Personnel Research						
Rating Scale						
Ford Program						
Qualitative Data						

U146269

DUDLEY KNOX LIBRARY - RESEARCH REPORTS
5 6853 01068091 1

~~U14626~~